





```
QY 3341 GTTCAGCTGACATTTCTACAAACACCACCACTCACCGATGTATTCGTTGCCAGTGGAA 2400
|||
DB 2341 GTTCAGCTGACATTTCTACAAACACCACCACTCACCGATGTATTCGTTGCCAGTGGAA 2400
|||
QY 2401 CATACAGAGCTGAATTTGGAAATAAATATGTGTTTCTTCCAGGAAATATCTAGACTG 2400
|||
DB 2401 CATACAGAGCTGAATTTGGAAATAAATATGTGTTTCTTCCAGGAAATATCTAGACTG 2400
|||
QY 2461 ACTTTGATGCTCCCAAAACATAACCCAGTGTAAAAACAGAGATGTGGAGGGAGCTGG 2520
|||
DB 2461 ACTTTGATGCTCCCAAAACATAACCCAGTGTAAAAACAGAGATGTGGAGGGAGCTGG 2520
|||
QY 2521 GAGATTTCAGTGGTACATTTGAATTCCTCCAACTTACCAGGCAATTAACCCAGCAACCG 2580
|||
DB 2521 GAGATTTCAGTGGTACATTTGAATTCCTCCAACTTACCAGGCAATTAACCCAGCAACCG 2580
|||
QY 2581 AGTGTAGTGGACCATCAACCCACCCCAAGGCGGATCTGATGCTGCTCCCTGAGA 2640
|||
DB 2581 AGTGTAGTGGACCATCAACCCACCCCAAGGCGGATCTGATGCTGCTCCCTGAGA 2640
|||
QY 2641 TCTTCTGCTCCATAGAGGAGCACTGTGGGACTATCTGCTGATGCGGAAACCTTCTCAT 2700
|||
DB 2641 TCTTCTGCTCCATAGAGGAGCACTGTGGGACTATCTGCTGATGCGGAAACCTTCTCAT 2700
|||
QY 2701 CCAATTCTGTGACAAATATGAACCTGCGAGACCTACGAACGCCCATCGCTTCACT 2760
|||
DB 2701 CCAATTCTGTGACAAATATGAACCTGCGAGACCTACGAACGCCCATCGCTTCACT 2760
|||
QY 2761 CCAAGTCAAAAGAGCTGTGGATTCAAGTCCAACTCAAGGCAACAGCGCTAGAGGT 2820
|||
DB 2761 CCAAGTCAAAAGAGCTGTGGATTCAAGTCCAACTCAAGGCAACAGCGCTAGAGGT 2820
|||
QY 2821 TCCAGTCCCATACGTGACATATGATGAGACTTACAGGAACTCAATTGAAGACATAGTTC 2880
|||
DB 2821 TCCAGTCCCATACGTGACATATGATGAGACTTACAGGAACTCAATTGAAGACATAGTTC 2880
|||
QY 2881 GAGATGCGAGCTCTATGATCTGAAACATCAGGAAATCTTAAGATAGAACTTA 2940
|||
DB 2881 GAGATGCGAGCTCTATGATCTGAAACATCAGGAAATCTTAAGATAGAACTTA 2940
|||
QY 2941 TCAAGGCTCTGTTTGTGCTCTGCGCCATCCCAAGTCAATTTCAAGTACAGGCGCCAG 3000
|||
DB 2941 TCAAGGCTCTGTTTGTGCTCTGCGCCATCCCAAGTCAATTTCAAGTACAGGCGCCAG 3000
|||
QY 3001 AGTCCGAGAGATGTTCCAGATCTGTCATCGATTGCTAGTTCACAAAGTTCAGGT 3060
|||
DB 3001 AGTCCGAGAGATGTTTCCAGATCTGTCATCGATTGCTAGTTCACAAAGTTCAGGT 3060
|||
QY 3061 TTTTGAGACCTTACAATGACTCAGCCACGTCGCCACTCAATACAAATCTCTGCTATAG 3120
|||
DB 3061 TTTTGAGACCTTACAATGACTCAGCCACGTCGCCACTCAATACAAATCTCTGCTATAG 3120
|||
QY 3121 GGTGTGGACAGAGCTGTTCTTCTTGTGATGTGTCAGCACAGTGGGTATTCCTGCTC 3180
|||
DB 3121 GGTGTGGACAGAGCTGTTCTTCTTGTGATGTGTCAGCACAGTGGGTATTCCTGCTC 3180
|||
QY 3181 CCGTATCAGTACTCATAGAGTTCAATTTTATAGATATACAGATATTTTGTAAAT 3240
|||
DB 3181 CCGTATCAGTACTCATAGAGTTCAATTTTATAGATATACAGATATTTTGTAAAT 3240
|||
QY 3241 GAACTTGGTCTTTCTTCCAGCATCGTGATAGACTGAGATGGCTTTGAGTGGCAT 3300
|||
DB 3241 GAACTTGGTCTTTCTTCCAGCATCGTGATAGACTGAGATGGCTTTGAGTGGCAT 3300
|||
QY 3301 CAGCTTCTACTGCTGGGGGATGCTTGGATAGATCACGGCTGGCTAGCTGGACT 3360
|||
DB 3301 CAGCTTCTACTGCTGGGGGATGCTTGGATAGATCACGGCTGGCTAGCTGGACT 3360
|||
QY 3361 TTGCTCAGCTTAGTGTAGACTCACTGCTCTCTGGGCTCTTACTCTCTCCTCAAGGAGTC 3420
|||
DB 3361 TTGCTCAGCTTAGTGTAGACTCACTGCTCTCTGGGCTCTTACTCTCTCCTCAAGGAGTC 3420
|||
QY 3421 TGTAGTGGAAAGAGGCGCCACAGAAATAGTGTATTCTTGAAACTTCAGACTTCTCTTAGC 3480
|||
```

```
DB 3421 TGTAGTGGAAAGAGGCGCCACAGAAATAGCTCTTATCTGAAACTTCAGCTTCTCTAGC 3480
|||
QY 3481 CCGGCCCTCTTAAGGAGCCCTCTGCACTGCTGTGCGAGGCTCTGACAGCAGAAACAGG 3540
|||
DB 3481 CCGGCCCTCTTAAGGAGCCCTCTGCACTGCTGTGCGAGGCTCTGACAGCAGAAACAGG 3540
|||
QY 3541 CAAGAGGCGAGGAGGAGGAGACCCCTGCAGGCTCCCTCCAGCCACCTTGAGACCTGGGAGG 3600
|||
DB 3541 CAAGAGGCGAGGAGGAGGAGACCCCTGCAGGCTCCCTCCAGCCACCTTGAGACCTGGGAGG 3600
|||
QY 3601 ACTCAGTCTCTCAACAGCTTCTCCAGCTGCTGTGATACAAAGTTTGATCCAGGACTTG 3660
|||
DB 3601 ACTCAGTCTCTCCAGAGCTTCTCCAGCTGCTGTGATACAAAGTTTGATCCAGGACTTG 3660
|||
QY 3661 AGTCTCAAGCAGTGTCTGCGGAAAGGAGGAGAAATAGAAATATAAATAAATACTA 3720
|||
DB 3661 AGTCTCAAGCAGTGTCTGCGGAAAGGAGGAGAAATAGAAATATAAATAAATACTA 3720
|||
QY 3721 AGCACTTCTGGAGACAT 3737
|||
DB 3721 AGCACTTCTGGAGACAT 3737
|||
```

## RESULT 2

```
US-10-177-293-44
; Sequence 44, Application US/10:77293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Ganraivaipu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Verrens, Maureen
; APPLICANT: Wyer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pasztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegül
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-018
; CURRENT APPLICATION NUMBER: US/10:177,293
; PRIOR FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,867
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,595
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows version 4.0
; SEQ ID NO 44
; LENGTH: 3737
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-293-44
```

Query Match: 100.0% Score 3737; DB 14; Length: 3737;







218 AGATGAGTGTGCCCAAGGGCTAGATGACTGCCATGCCGAGCCCTGTGTGACAAACACC 277  
221 AGATGAGTGTCCGCAAGGGCTAGATGACTGCCATGCCGAGCCCTGTGTGACAAACACC 280  
278 CACCTCCCTACAAGTGTCTCTGCAAGCCCTGGCTACCAAGGGGAAGGCAGGAGTGTGAGGA 337  
281 CACCTCCCTACAAGTGTCTCTGCAAGCCCTGGCTACCAAGGGGAAGGCAGGAGTGTGAGGA 340  
338 CATCGATGAATGTGGAATGAGCTCAATGGAGGGCTGTGTCATGACACTGTTTGAATATTC 337  
341 CATCGATGAATGTGGAATGAGCTCAATGGAGGGCTGTGTCATGACACTGTTTGAATATTC 400  
398 AGGCAATTTATCGTTGCACATGTTTGTGATGGCTTCATGTTGGCTCATGACGGTCAATATTC 457  
401 AGGCAATTTATCGTTGCACATGTTTGTGATGGCTTCATGTTGGCTCATGACGGTCAATATTC 460  
458 TCATTGATGTGACAGTGTGCTGCGAGAGCAATGGCGGTGTCAGCATACCTGTGTCAACGT 517  
461 TCATTGATGTGACAGTGTGCTGCGAGAGCAATGGCGGTGTCAGCATACCTGTGTCAACGT 520  
518 CATGGGAGCTATCAGTGTGCTGCGAGAGGGGCTTTTCTGAGTGACAAATCAGCACAC 577  
521 CATGGGAGCTATCAGTGTGCTGCGAGAGGGGCTTTTCTGAGTGACAAATCAGCACAC 580  
578 CTGCATTTACCGCTCGGAAGAGGGCTGAGCTGCATGAATAAGGATCACGGCTGTAGTCA 637  
581 CTGCATTTACCGCTCGGAAGAGGGCTGAGCTGCATGAATAAGGATCACGGCTGTAGTCA 640  
638 CATCTGGAAGAGGCCCAAGGGCAGGCTGCGCTGTGAGTGAGGCTCGTTTGTAGCT 697  
641 CATCTGGAAGAGGGCCCAAGGGCAGGCTGCGCTGTGAGTGAGGCTCGTTTGTAGCT 700  
698 GGCCAAAGAACAGAGAGACTGCATCTTGACTGTAAACATGGGGAACGGTGGTGCACGA 757  
702 GGCCAAAGAACAGAGAGACTGCATCTTGACTGTAAACATGGGGAACGGTGGTGCACGA 760  
758 CTCTGTGACGATACAGCCGATGAGCCAGAGTGCAGCTGCCATCCACAGTACAAGATGCA 817  
761 CTCTGTGACGATACAGCCGATGAGCCAGAGTGCAGCTGCCATCCACAGTACAAGATGCA 820  
818 CACAGATGGGAGGAGCTGCCCTTGAGTCAGAGGACACTGTCTGGAGGTGACAGAGACAA 977  
821 CACAGATGGGAGGAGCTGCCCTTGAGTCAGAGGACACTGTCTGGAGGTGACAGAGACAA 980  
878 CACCACATCAGTGTGATGGGATAAACGGGTGAAACGGCGCTGCTCAAGGAACGTCG 937  
881 CACCACATCAGTGTGATGGGATAAACGGGTGAAACGGCGCTGCTCAAGGAACGTCG 940  
938 TGCTGTCAACAAATGGAGGCTGTGACCGACCTGTGAAGGATACCTPGACAGGTGTCCACTG 997  
941 TGCTGTCAACAAATGGAGGCTGTGACCGACCTGTGAAGGATACCTPGACAGGTGTCCACTG 1000  
998 CAGTTGCTCTTGTGATTTCACTCTCCAGTTGGATGGGAACACATGTAAGATATTGATGA 1057  
1001 CAGTTGCTCTTGTGATTTCACTCTCCAGTTGGATGGGAACACATGTAAGATATTGATGA 1060  
1058 GTGCCAGACCCGCAATGGAGGCTGTGATTCATTTCTGCAAAAACATCGTGGCGAGTTTGA 1117  
1061 GTGCCAGACCCGCAATGGAGGCTGTGATTCATTTCTGCAAAAACATCGTGGCGAGTTTGA 1120  
1118 CTGGCGCTGCAAGAAAGGATTTAAATTAATAACAGATGAGAAGTCTTGCACAAATGTGGA 1177  
1121 CTGGCGCTGCAAGAAAGGATTTAAATTAATAACAGATGAGAAGTCTTGCACAAATGTGGA 1180  
1178 TGAGTGTCTTTGTGATAGGACTGTGACACAGAGCTGCATCAACCCCTGGCAATTTGC 1237  
1181 TGAGTGTCTTTGTGATAGGACTGTGACACAGAGCTGCATCAACCCCTGGCAATTTGC 1240  
1238 TTGTGCTTGCACCGAGGATACCCCTGTATGGCTTCAACCCACTGTGGAGACACCAATGA 1297  
1241 TTGTGCTTGCACCGAGGATACCCCTGTATGGCTTCAACCCACTGTGGAGACACCAATGA 1300  
1298 GTGCAGCATCAACCAAGAGGCTGTGACGAGGCTGTGTGTAACACAGTGGGCGACTATGA 1357

1301 GTGCAGCATCAACCAAGAGGCTGTGACAGGCTGTGTGTGAACACAGTGGGCGACTATGA 1360  
1358 ATGCCAGTGTCCACCCCTGGGTACAAGCTCACTGGAAATAAAAAGACTGTGTGTGAAGTGA 1417  
1361 ATGCCAGTGTCCACCCCTGGGTACAAGCTCACTGGAAATAAAAAGACTGTGTGTGAAGTGA 1420  
1418 GGGGTCTCTGCCCAACAAGTGTGTACCCCGGTGTGTCCTGGACCTCGGTAAAGTGGTGG 1477  
1421 GGGGTCTCTGCCCAACAAGTGTGTACCCCGGTGTGTCCTGGACCTCGGTAAAGTGGTGG 1480  
1478 AGGAGACGGTGTCTCTCCTCAGATGTCACTTGGCATTACCTCTCTTCCAGATGTCAACAC 1537  
1491 AGGAGACGGTGTCTCTCCTCAGATGTCACTTGGCATTACCTCTCTTCCAGATGTCAACAC 1540  
1538 CATCAGGACCAAGTGTAACTTTAAGCTAAATAAGGCAAGTGTAGTTTGAATAAATGCTGA 1597  
1541 CATCAGGACCAAGTGTAACTTTAAGCTAAATAAGGCAAGTGTAGTTTGAATAAATGCTGA 1600  
1598 GCTGTCTCCCGAGGGTCTGCGACACGACTACCAAGAGGACACACTCAGTAAAGAGAG 1657  
1601 GCTGTCTCCCGAGGGTCTGCGACACGACTACCAAGAGGACACACTCAGTAAAGAGAG 1660  
1658 CTCTCGCTACGTAAACCTTACATGACGCTCTGGCAAGCAAGTCCAGAGGAGCCCTGGCG 1717  
1661 CTCTCGCTACGTAAACCTTACATGACGCTCTGGCAAGCAAGTCCAGAGGAGCCCTGGCG 1720  
1718 ACCAAGACCCCTAAGGAATAGTTTATCACTGTTTGAAGTGTGAGCTTGAACCTTAACAAA 1777  
1721 ACCAAGACCCCTAAGGAATAGTTTATCACTGTTTGAAGTGTGAGCTTGAACCTTAACAAA 1780  
1728 GAGGTGTACAGCTCTTGTGACCTGAGCTGCATCTGAAGCGAACCGAGAGGCGCTCG 1837  
1781 GAGGTGTACAGCTCTTGTGACCTGAGCTGCATCTGAAGCGAACCGAGAGGCGCTCG 1840  
1838 TAAAGCCATCCCAACGCTCAGAAAGCGCTCCACAGGAGGAGTTCACCTCCAGCTCTC 1897  
1841 TAAAGCCATCCCAACGCTCAGAAAGCGCTCCACAGGAGGAGTTCACCTCCAGCTCTC 1900  
1898 AGGCATGACCTCGACGTGGCTTAAAGCCCTCCAGAGACATCTGAACCGCAGGAGAGTTC 1957  
1901 AGGCATGACCTCGACGTGGCTTAAAGCCCTCCAGAGACATCTGAACCGCAGGAGAGTTC 1960  
1958 CTGTGAGTGGGCCAGGGTCAATGCAAGAAACCAATGTGTGAGTCCAGGGCTGGACCTTA 2017  
1961 CTGTGAGTGGGCCAGGGTCAATGCAAGAAACCAATGTGTGAGTCCAGGGCTGGACCTTA 2020  
2018 TTATGATGGAGCAACGAAACGCTGCAATTTATGTCCAAATGGAACCTTCCAAATGAGGA 2077  
2021 TTATGATGGAGCAACGAAACGCTGCAATTTATGTCCAAATGGAACCTTCCAAATGAGGA 2080  
2078 AGGACAAATGACTGTGAACCATGCCAAGACCCAGAGAAATCTGGGGCCCTGAAGACCCC 2137  
2081 AGGACAAATGACTGTGAACCATGCCAAGACCCAGAGAAATCTGGGGCCCTGAAGACCCC 2140  
2138 AGAAGCTTGGAAATATCTGTAATGTGAGGCTGTGTCAACCTGTGTGAATATCTTCGAGA 2197  
2141 AGAAGCTTGGAAATATCTGTAATGTGAGGCTGTGTCAACCTGTGTGAATATCTTCGAGA 2200  
2198 TGGCTTTGACCTTCCAGCTGTGTGCCCTGGGACGTTCCAGCTGGAAGTGTGCGAAC 2257  
2201 TGGCTTTGACCTTCCAGCTGTGTGCCCTGGGACGTTCCAGCTGGAAGTGTGCGAAC 2260  
2258 TTCTGCTCTCCCTGTGGAGAGGCTTGCACCAACAAACATCAGGGAGCTACTTCTTTCA 2317  
2261 TTCTGCTCTCCCTGTGGAGAGGCTTGCACCAACAAACATCAGGGAGCTACTTCTTTCA 2320  
2318 GGAAGTGTGAACCAAGTGTGAATGTTCACCTGGACATTTCTACACACACACACCTCAACG 2377  
2321 GGAAGTGTGAACCAAGTGTGAATGTTCACCTGGACATTTCTACACACACACACCTCAACG 2380  
2378 ATGTATTCTTCTGCCCAGTGGGAACATAACAGCTTGAATTTGGAAAAATAAATGTGTTC 2437



Db 401 AGGCAATTATCGTTGCACTGTGTTTGTATGGCTTCAATGTTGGCTCATGACGGTCAATAATTG 460  
QY 458 TCTTGATGTGGACGAGTGTCTGGAGAAACAATGGGGCTGCCAGCATACCTGTGTCAACGTT 517  
Db 461 TCTTGATGTGGACGAGTGTCTGGAGAAACAATGGGGCTGCCAGCATACCTGTGTCAACGTT 520  
QY 518 CATGGGGAGCTATGAGTGTCTGTCAAGAGAGGGTTTTTCTGTAGTGACAAATCAGACAC 577  
Db 521 CATGGGGAGCTATGAGTGTCTGTCAAGAGAGGGTTTTTCTGTAGTGACAAATCAGACAC 580  
QY 578 CTGCATTACCCCTCGGAAGAGGGCTGTAGCTGTCAATGAATAGGATCACGGCTGTAGTCA 637  
Db 581 CTGCATTACCCCTCGGAAGAGGGCTGTAGCTGTCAATGAATAGGATCACGGCTGTAGTCA 640  
QY 638 CATCTGCAAGAGGCCCAAGGGGAGGGTGTCTGTGAGTGTGAGGCTGTGTTTGTAGCT 697  
Db 641 CATCTGCAAGAGGCCCAAGGGGAGGGTGTCTGTGAGTGTGAGGCTGTGTTTGTAGCT 700  
QY 698 GGCCAAAGAACAGAGAGACTGTGATCTTGACCTGTAAACCATGGAAACGGTGGGCCAGCA 757  
Db 701 GGCCAAAGAACAGAGAGACTGTGATCTTGACCTGTAAACCATGGAAACGGTGGGCCAGCA 760  
QY 758 CTCCTGTGACGATACAGCCGATGGCCAGAGTGCAGCTGCCATCCACAGTACAAAGATSCA 817  
Db 761 CTCCTGTGACGATACAGCCGATGGCCAGAGTGCAGCTGCCATCCACAGTACAAAGATSCA 820  
QY 818 CACAGATGGGAGAGTGCCTTGAGGGAGAGACACTGTCTGTGAGAGTGACAGAGACAA 877  
Db 821 CACAGATGGGAGAGTGCCTTGAGGGAGAGACACTGTCTGTGAGAGTGACAGAGACAA 880  
QY 878 CACCACATCAGTGGTGGATGGGGATAAAGCGAGTGAACCGGGCTGTCTCATGGAAACGTG 937  
Db 881 CACCACATCAGTGGTGGATGGGGATAAAGCGAGTGAACCGGGCTGTCTCATGGAAACGTG 940  
QY 938 TGCCTGTCAACATGAGAGCTGTGACCGCACCTGTGAGGATCTTCGACAGTGTCTCACTG 997  
Db 941 TGCCTGTCAACATGAGAGCTGTGACCGCACCTGTGAGGATCTTCGACAGTGTCTCACTG 1000  
QY 998 CAGTGTCTCTGTGATTCACCTCTCCAGTTGGATGGGAAGACATGTAAAGATATTGATGA 1057  
Db 1001 CAGTGTCTCTGTGATTCACCTCTCCAGTTGGATGGGAAGACATGTAAAGATATTGATGA 1060  
QY 1058 GTGCCAGACCCGCAATGGAGTGTGATCATTTCTGCAAAAACAATCGTGGCAGTTTTGA 1117  
Db 1061 GTGCCAGACCCGCAATGGAGTGTGATCATTTCTGCAAAAACAATCGTGGCAGTTTTGA 1120  
QY 1118 CTGGGCTGCAAGAAGAGATTTAAATTAATTAACAGATCAGAAAGTCTTCCCAAGATGTGA 1177  
Db 1121 CTGGGCTGCAAGAAGAGATTTAAATTAATTAACAGATCAGAAAGTCTTCCCAAGATGTGA 1180  
QY 1178 TGAGTGTCTTTTGGATAGGACTGTGACACAGCTGCAATCAACCACTGTGGAGACCCATGA 1237  
Db 1181 TGAGTGTCTTTTGGATAGGACTGTGACACAGCTGCAATCAACCACTGTGGAGACCCATGA 1240  
QY 1238 TTGTGCTTGCACCGAGGTACACCTGTATGGCTTACCCACTGTGGAGACCCATGA 1297  
Db 1241 TTGTGCTTGCACCGAGGTACACCTGTATGGCTTACCCACTGTGGAGACCCATGA 1300  
QY 1298 GTGCAGCATCAACAACGAGGCTGTGAGCAGGTCTGTGTGAACACAGTGGGAGCTATGA 1357  
Db 1301 GTGCAGCATCAACAACGAGGCTGTGAGCAGGTCTGTGTGAACACAGTGGGAGCTATGA 1360  
QY 1358 ATGCCAGTGGCACCCTGGGTACAGCTGCATGGAAATAAAGAAAGTGTGTGGAAAGTGA 1417  
Db 1361 ATGCCAGTGGCACCCTGGGTACAGCTGCATGGAAATAAAGAAAGTGTGTGGAAAGTGA 1420  
QY 1418 GGGGCTCTGCCCAAGTGTGACCCCGTGTGTCTGCTGACCTGCGGTAGAGGTGGTGG 1477  
Db 1421 GGGGCTCTGCCCAAGTGTGACCCCGTGTGTCTGCTGACCTGCGGTAGAGGTGGTGG 1480  
QY 1478 AGGAGACGGGTGCTTCTCAGATGTCACTCTGGCATTTCACTCTCTTCAGATGTCAACAC 1537  
Db 1481 AGGAGACGGGTGCTTCTCAGATGTCACTCTGGCATTTCACTCTCTTCAGATGTCAACAC 1540

QY 1538 CATCAGGCAAGTGTAAACCTTTAAAGCTAAATGAAGGCAAGTGTAGTTTGAAAAATGCTGA 1597  
Db 1541 CACCAGGCAAGTGTAAACCTTTAAAGCTAAATGAAGGCAAGTGTAGTTTGAAAAATGCTGA 1600  
QY 1598 GCTGTTTCCGAGAGGTCTGGACCCAGCACTACCAAGAAAGCAAGCTCAATAAGAGAG 1657  
Db 1601 GCTGTTTCCGAGAGGTCTGGACCCAGCACTACCAAGAAAGCAAGCTCAATAAGAGAG 1660  
QY 1658 CTTCCGCTACGCTTAACCTTACATGAGCTGTGSSAAGCAAGTCCAGGAGCCCTGSCCG 1717  
Db 1661 CTTCCGCTACGCTTAACCTTACATGAGCTGTGSSAAGCAAGTCCAGGAGCCCTGSCCG 1720  
QY 1718 ACCAAGCAACCTTAAGGAAATGTTTATCATCTGTTGAGTTTGAAGCTTGAATTAACCAAA 1777  
Db 1721 ACCAAGCAACCTTAAGGAAATGTTTATCATCTGTTGAGTTTGAAGCTTGAATTAACCAAA 1780  
QY 1778 GGAGGTGACAGCTTCTTGTGACCTGAGCTGATCGTAAAGCGAACCGAGAGCGGTCCG 1837  
Db 1781 GGAGGTGACAGCTTCTTGTGACCTGAGCTGATCGTAAAGCGAACCGAGAGCGGTCCG 1840  
QY 1838 TAAAGCCATCCGACAGCTCAGAAAGGCGCTCCAGAGGAGCAGTTTCACTCCAGCTTC 1897  
Db 1841 TAAAGCCATCCGACAGCTCAGAAAGGCGCTCCAGAGGAGCAGTTTCACTCCAGCTTC 1900  
QY 1898 AGGCAATGAGCTTGAAGCTTGAAGAGCGCTTGAAGAGCGCTTGAAGAGCGCGAGCTC 1957  
Db 1901 AGGCAATGAGCTTGAAGCTTGAAGAGCGCTTGAAGAGCGCTTGAAGAGCGCGAGCTC 1960  
QY 1958 CTGTGGATGGGCGGAGGCTTATGCAAGAAACCAATGTGTGAGTTGCAAGGCTGGAGCTA 2017  
Db 1961 CTGTGGATGGGCGGAGGCTTATGCAAGAAACCAATGTGTGAGTTGCAAGGCTGGAGCTA 2020  
QY 2018 TATGATGGAGCAGAGAACCTGCTGATTTATGTCCAAATGGAACTTCCAAATGAGGA 2077  
Db 2021 TATGATGGAGCAGAGAACCTGCTGATTTATGTCCAAATGGAACTTCCAAATGAGGA 2080  
QY 2078 AGGCAAAATGACTTGTGAACCATGCCAAGACAGGAAATTCGGGGCTTGAAGACCCC 2137  
Db 2081 AGGCAAAATGACTTGTGAACCATGCCAAGACAGGAAATTCGGGGCTTGAAGACCCC 2140  
QY 2138 AGAAGCTTGAATATGTCTGAATGTGAGGTCTGTGTCAACCTGTGTGAATATTCTGAGA 2197  
Db 2141 AGAAGCTTGAATATGTCTGAATGTGAGGTCTGTGTCAACCTGTGTGAATATTCTGAGA 2200  
QY 2198 TGGCTTTCACCTTGCACAGCTGTGCGCTGGGACCTTCCAGCTCAAGCTGGTCCAAAC 2257  
Db 2201 TGGCTTTCACCTTGCACAGCTGTGCGCTGGGACCTTCCAGCTCAAGCTGGTCCAAAC 2260  
QY 2258 TTCTGTCTTCCCTGTGGAGAGGCTTGCACAACCAATCAGGAGCTACTTCTTTCA 2317  
Db 2261 TTCTGTCTTCCCTGTGGAGAGGCTTGCACAACCAATCAGGAGCTACTTCTTTCA 2320  
QY 2318 GAGCTGTCAAAACAGAGTCAATGTTCACCTGGACATTTCTACAACACCACTACACCG 2377  
Db 2321 GAGCTGTCAAAACAGAGTCAATGTTCACCTGGACATTTCTACAACACCACTACACCG 2380  
QY 2378 ATGTATTCTGTTGCCAGTGGAAACATACCAAGCTTGAATTTGGAAAAAATTAATTTGTTTC 2437  
Db 2381 ATGTATTCTGTTGCCAGTGGAAACATACCAAGCTTGAATTTGGAAAAAATTAATTTGTTTC 2440  
QY 2438 TTGCCAGGAATATACAGCTGACTTGTGATGGCTCCAGAAACATACCCAGTGTAAAAA 2497  
Db 2441 TTGCCAGGAATATACAGCTGACTTGTGATGGCTCCAGAAACATACCCAGTGTAAAAA 2500  
QY 2498 CAGAAGATGTGGAGGAGCTGGAGATTTCACTGGGTACATTTGAATCCCAAACTACCC 2557  
Db 2501 CAGAAGATGTGGAGGAGCTGGAGATTTCACTGGGTACATTTGAATCCCAAACTACCC 2560  
QY 2558 AGGCAATTACCAAGCCCAACAGCGAGTGTAGTGAACCATCAACCCAGCCCGGAGCGCG 2617  
Db 2561 AGGCAATTACCAAGCCCAACAGCGAGTGTAGTGAACCATCAACCCAGCCCGGAGCGCG 2620



QY 917 GCGGCTGCTCATGGAAAGTGTGTGTCTCAAAATGAGGCTGTGACCGCACCTCTGAAGGA 976  
DB |||||  
QY 556 GCGGCTGCTCATGGAAAGTGTGTGTCTCAAAATGAGGCTGTGACCGCACCTCTGAAGGA 615  
DB |||||  
QY 977 TACTTCGACAGGTGTCCACTGCGAGTGTCTCTGTGGATTCACCTCTCCAGTTGGATGGGAA 1036  
DB |||||  
QY 616 TACTTCGACAGGTGTCCACTGCGAGTGTCTCTGTGGATTCACCTCTCCAGTTGGATGGGAA 675  
DB |||||  
QY 1037 GACATGTAAAGATATTGATGAGTGCCAGACCGGCAATGGAGGTTGTGATCAATTTCTGGAA 1096  
DB |||||  
QY 676 GACATGTAAAGATATTGATGAGTGCCAGACCGGCAATGGAGGTTGTGATCAATTTCTGGAA 735  
DB |||||  
QY 1097 AAACATCTGGGCGAGTTTGTGCTGGGCTGCAAGAAAGGATTTAAATTTATTAACAGATGA 1156  
DB |||||  
QY 736 AAACATCTGGGCGAGTTTGTGCTGGGCTGCMAGAAAGGATTTAAATTTATTAACAGATGA 795  
DB |||||  
QY 1157 GAAGTCTTTGCCAAGATGTGGATGAGTGTCTTTTGGATAGGACCTGTGACCAACAGCTGCAT 1216  
DB |||||  
QY 796 GAAGTCTTTGCCAAGATGTGGATGAGTGTCTTTTGGATAGGACCTGTGACCAACAGCTGCAT 855  
DB |||||  
QY 1217 CAACACCTGTGCATTTGCTTGTGTGTGCTTGCNACCGAGGTTACACCTGTATGCTTCAC 1276  
DB |||||  
QY 856 CAACACCTGTGCATTTGCTTGTGTGCTTGTGCAACCGAGGGTACACCTGTATGCTTCAC 915  
DB |||||  
QY 1277 CCACCTGTGGAGACACCAATGAGTGCAGATCAACAACGAGAGGCTGTGACGAGGTCTGTGT 1336  
DB |||||  
QY 916 CCACCTGTGGAGACACCAATGAGTGCAGATCAACAACGAGAGGCTGTGACGAGGTCTGTGT 975  
DB |||||  
QY 1337 GAACACAGTGGGACGCTATGAATGCCAGTGCACACCTGGGTACAAGCTCCACTGSAATAA 1396  
DB |||||  
QY 976 GAACACAGTGGGACGCTATGAATGCCAGTGCACACCTGGGTACAAGCTCCACTGSAATAA 1035  
DB |||||  
QY 1397 AAAGACCTGTGGAAGTCAAGGGCTCTGTGCCACCAAGTGTGTACCCCGTGTCTCCT 1456  
DB |||||  
QY 1036 AAAGACCTGTGGAAGTCAAGGGCTCTGTGCCACCAAGTGTGTACCCCGTGTGTCTCCT 1095  
DB |||||  
QY 1457 GCACCTGCGGTAAAGTGTGGAGGAGAGGGGTGTCTTCTCAGATGTCACTCTGGCAATTCA 1516  
DB |||||  
QY 1096 GCACCTGCGGTAAAGTGTGGAGGAGAGGGGTGTCTTCTCAGATGTCACTCTGGCAATTCA 1155  
DB |||||  
QY 1517 CTTCTCTTCA----- 1526  
DB |||||  
QY 1156 CTTCTCTCAGGACTGCAAGGGGCTTACTGTCTCACTGTGGCTCTTCTCTCTCTCAG 1215  
DB |||||  
QY 1527 -----GATGTCAACCACTCAGGACAAG 1549  
DB |||||  
QY 1216 GAACAAACAAACAAATCAATGACTCTGTCTTTTGGGATGTCAACCACTCAGGACAAG 1275  
DB |||||  
QY 1550 TGTAACTTTAAGCTAAATGAAGGCAAGTGTAGTTTGAAATAATGCTGAGCTGTTTCCCGA 1609  
DB |||||  
QY 1276 TGTAACTTTAAGCTAAATGAAGGCAAGTGTAGTTTGAAATAATGCTGAGCTGTTTCCCGA 1335  
DB |||||  
QY 1610 GGGTCTGGACAGCACTACAGAGAAGCAGAGCTCAGTAAAGAGAGCTTCCGCTAGCT 1669  
DB |||||  
QY 1336 GGGTCTGGACAGCACTACAGAGAAGCAGAGCTCAGTAAAGAGAGAGCTTCCGCTAGCT 1395  
DB |||||  
QY 1670 AAACCTTACATGACGCTCTGGCAAGCAAGTCCCGAGGAGCCCTGCGCCACCAAGCACCCC 1729  
DB |||||  
QY 1396 AAACCTTACATGACGCTCTGGCAAGCAAGTCCCGAGGAGCCCTGCGCCACCAAGCACCCC 1455  
DB |||||  
QY 1730 TAAGGAATGTTTATCACTGTTGAGTTGAGTTGAAACTAACCAAAAGGAGGTGACAGC 1789  
DB |||||  
QY 1456 TAAGGAATGTTTATCACTGTTGAGTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1515  
DB |||||  
QY 1790 TTTCTGTGACCTGAGCTGATCGTAAAGCGAACCGAGAGCGGCTCCGTTAAAGCATCCG 1849  
DB |||||  
QY 1516 TTTCTGTGACCTGAGCTGATCGTAAAGCGAACCGAGAGCGGCTCCGTTAAAGCATCCG 1575  
DB |||||  
QY 1850 CACGCTCAGAAAGGCGCTCCACAGGAGCAGTTTCACTTCCAGCTCTCAGGATCAAGCT 1909  
DB |||||  
QY 1576 CACGCTCAGAAAGGCGCTCCACAGGAGCAGTTTCACTTCCAGCTCTCAGGATCAAGCT 1635  
DB |||||

QY 1910 CGAGGTGGCTAAAGAGCTCCCGAGAACATCTTGAAAGCCAGGCGAGAGTCTCTGTGGAGTGGG 1969  
DB |||||  
QY 1636 CGAGGTGGCTAAAGAGCTCCCGAGAACATCTTGAAAGCCAGGCGAGAGTCTCTGTGGAGTGGG 1695  
DB |||||  
QY 1970 CCAGGTCTATGAGAAACCAATGTGTGAGTGTGAGGCTGGAGCTTATATGATGGAGC 2029  
DB |||||  
QY 1696 CCAGGTCTATGAGAAACCAATGTGTGAGTGTGAGGCTGGAGCTTATATGATGGAGC 1755  
DB |||||  
QY 2030 ACAGAGACCTGATTTATGTCGAATGGAACTTCCAAATGAGGAGGAGGAGCAATGAC 2089  
DB |||||  
QY 1756 ACAGAGACCTGATTTATGTCGAATGGAACTTCCAAATGAGGAGGAGGAGCAATGAC 1815  
DB |||||  
QY 2090 TTGTGAACCATGCCAAGACCAAGAAATTTCTGGGCGCTTGAAGACCCAGAAAGCTTGGAA 2149  
DB |||||  
QY 1816 TTGTGAACCATGCCAAGACCAAGAAATTTCTGGGCGCTTGAAGACCCAGAAAGCTTGGAA 1875  
DB |||||  
QY 2150 TATGTTGAATGTGAGGTCTGTGTCAACCTGGTGAATATTTCTGAGATGGCTTGGAC 2209  
DB |||||  
QY 1876 TATGTTGAATGTGAGGTCTGTGTCAACCTGGTGAATATTTCTGAGATGGCTTGGAC 1935  
DB |||||  
QY 2210 TTGCCAGCTGTGTGCGCTGGGACCTTCCAGCTGAGCTGTGAGACTTCTGCTTCC 2269  
DB |||||  
QY 1936 TTGCCAGCTGTGTGCGCTGGGACCTTCCAGCTGAGCTGTGAGACTTCTGCTTCC 1995  
DB |||||  
QY 2270 CTGTGAGGAGGCTTGGCCACCAACATCAGGAGCTTCTTCTTCCAGGACTGTGAAAC 2323  
DB |||||  
QY 1996 CTGTGAGGAGGCTTGGCCACCAACATCAGGAGCTTCTTCTTCCAGGACTGTGAAAC 2055  
DB |||||  
QY 2330 CAGAGTTCAATGTTCACCTGGACATTTCTAACACCACTCAGGAGTGTATGCTGG 2399  
DB |||||  
QY 2056 CAGAGTTCAATGTTCACCTGGACATTTCTAACACCACTCAGGAGTGTATGCTGG 2115  
DB |||||  
QY 2390 CCAGTGTGGAACATACCAAGCTGAAATTTGGAATAATTAATGCTTCTGCGCAGGAA 2449  
DB |||||  
QY 2116 CCAGTGTGGAACATACCAAGCTGAAATTTGGAATAATTAATGCTTCTGCGCAGGAA 2175  
DB |||||  
QY 2450 TACTAGACTGACTTTGATGGCTCCACAAACATAACCCAGTGTAAACACAGAGATGG 2509  
DB |||||  
QY 2176 TACTAGACTGACTTTGATGGCTCCACAAACATAACCCAGTGTAAACACAGAGATGG 2235  
DB |||||  
QY 2510 AGGGAGCTGGAGATTTCACTGGGTAGATTAATGCTTCTGCGCAGGAA 2569  
DB |||||  
QY 2236 AGGGAGCTGGAGATTTCACTGGGTAGATTAATGCTTCTGCGCAGGAA 2295  
DB |||||  
QY 2570 AGCCACACGAGTGTACCTGGACATCAACCCACCCAGGCGGCTCTGATCT 2629  
DB |||||  
QY 2296 AGCCACACGAGTGTACCTGGACATCAACCCACCCAGGCGGCTCTGATCT 2355  
DB |||||  
QY 2630 GGTCTCTGAGATTTCTCTGCGCAGAGGAGCTGTGGGACTATCTGTGATGGGAA 2689  
DB |||||  
QY 2356 GGTCTCTGAGATTTCTCTGCGCAGAGGAGCTGTGGGACTATCTGTGATGGGAA 2415  
DB |||||  
QY 2690 AACCTCTTCACTCAATTTCTGTGACACATATGAACCTGCGCAGACTCAGAACGCCAT 2749  
DB |||||  
QY 2416 AACCTCTTCACTCAATTTCTGTGACACATATGAACCTGCGCAGACTCAGAACGCCAT 2475  
DB |||||  
QY 2750 CGCTTTCACTCCAGTCAAGAAAGCTGTGATTTAGTTTCAAGTCCCAATGAAGGAAAG 2809  
DB |||||  
QY 2476 CGCTTTCACTCCAGTCAAGAAAGCTGTGATTTAGTTTCAAGTCCCAATGAAGGAAAG 2535  
DB |||||  
QY 2810 CGTATAGAGGTTCCAGGTCCTATGAGACTATGATGAGGACTACCGAGAACTCATGCA 2869  
DB |||||  
QY 2536 CGTATAGAGGTTCCAGGTCCTATGAGACTATGATGAGGACTACCGAGAACTCATGCA 2595  
DB |||||  
QY 2870 AGACATAGTTGAGAGTGGCAGGCTCTATGATCTGAGACCATCAGGAATACTTAAGA 2929  
DB |||||  
QY 2596 AGACATAGTTGAGAGTGGCAGGCTCTATGATCTGAGAACTCAGGAATACTTAAGA 2655  
DB |||||  
QY 2930 TAAGAACTTATCAAGGCTCTGTTTGTGCTTGGGCTTCCCGCAACTATTTCAAGTA 2989  
DB |||||  
QY 2656 TAAGAACTTATCAAGGCTCTGTTTGTGCTTGGGCTTCCCGCAACTATTTCAAGTA 2715  
DB |||||  
QY 2990 CACAGTCCAGGAGTCCCGAGAGATGTTTCCAGATGCTTTCATCGGATTTCTACGTTCCAA 3049  
DB |||||

DB 2726 CACAGCCAGGAGTCCCGAGAGATGTTCCAGATCGTTCATCCGATGCTACGTTCCAA 2775  
QY 3050 AGTGTCAGGTTTTCAGACCTTACAAATGACTCAGCCACCGTGCCACTCAATACAAATG 3109  
DB 2776 AGTGTCAGGTTTTCAGACCTTACAAATGACTCAGCCACCGTGCCACTCAATACAAATG 2835  
QY 3110 TTCTGCTATAGGTTGGTGGGACAGAGCTGCTTCTCTGCTATGTCAGCACAGTCGGGT 3169  
DB 2836 TTCTGCTATAGGTTGGTGGGACAGAGCTGCTTCTCTGCTATGTCAGCACAGTCGGGT 2895  
QY 3170 ATTGCTGCTCCCGTATCAGTGACTCATTAGAGTTCATTTTATAGATAATACAGATAT 3229  
DB 2896 ATTGCTGCTCCCGTATCAGTGACTCATTAGAGTTCATTTTATAGATAATACAGATAT 2955  
QY 3230 TTGTGTAATGAA 3243  
DB 2956 TTGTGTAATGAA 2969

## RESULT 6

US-09-898-570-11/c  
; Sequence 11, Application US/09898570  
; Patent No. US20020123612A1  
; GENERAL INFORMATION:  
; APPLICANT: GERLACH, VALERIE L.  
; APPLICANT: ELLERMAN, KAREN  
; APPLICANT: MACDOUGALL, JOHN R.  
; APPLICANT: SMITHSON, GLENNDA  
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND  
; TITLE OF INVENTION: METHODS OF USING THE SAME  
; FILE REFERENCE: 15966-776CIP

; CURRENT APPLICATION NUMBER: US/09/898,570  
; CURRENT FILING DATE: 2001-07-03  
; PRIOR APPLICATION NUMBER: 60/198,293  
; PRIOR FILING DATE: 2000-04-19  
; PRIOR APPLICATION NUMBER: 60/198,645  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: 60/210,809  
; PRIOR FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: 60/199,476  
; PRIOR FILING DATE: 2000-04-26  
; PRIOR APPLICATION NUMBER: 60/200,025  
; PRIOR FILING DATE: 2000-04-26  
; PRIOR APPLICATION NUMBER: 60/224,610  
; PRIOR FILING DATE: 2000-08-11  
; PRIOR APPLICATION NUMBER: 60/200,024  
; PRIOR FILING DATE: 2000-04-26  
; PRIOR APPLICATION NUMBER: 60/199,880  
; PRIOR FILING DATE: 2000-04-26  
; PRIOR APPLICATION NUMBER: 60/218,591  
; PRIOR FILING DATE: 2000-07-17  
; PRIOR APPLICATION NUMBER: 60/271,814  
; PRIOR FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: 60/215,855  
; PRIOR FILING DATE: 2000-07-03  
; PRIOR APPLICATION NUMBER: 09/839,446  
; PRIOR FILING DATE: 2001-04-19  
; NUMBER OF SEQ ID NOS: 58  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 11  
; LENGTH: 4821  
; TYPE: DNA  
; ORGANISM: Unknown Organism  
; FEATURE:  
; OTHER INFORMATION: Description of Unknown. Organism: POLYX  
; OTHER INFORMATION: 297832\_B.0.707  
US-09-898-570-11

Query Match 28.6%; Score 1069.2; DB 10; Length 4821;  
Best Local Similarity 63.1%; Pred. No. 0;  
Matches 1836; Conservative 0; Mismatches 993; Indels 8; Gaps 9;

QY 196 CCGCGGSCCCAGGAGATGTAGATCAGTGTGCCCCAAGGGCTAGATGACTGCCATGCCG 255  
DB 4412 CCCCTCTGCCCTGCCAGATGTGATGTGTGGAGGGGACTGACAATGCCACATCG 4353  
QY 256 ACGCCCTGTGTGAGACACACCCACCTCCTACAGTGTCTCTCCAGCCTGGCTACCAAG 315  
DB 4352 ATGTAATCTGCCAGAACCCCGAGGTCTACAGTGTCTCTCCAGTGTGGCTTACAG 4293  
QY 316 GGGAGGACGAGCTGTGAGGACATCCGATGAATGTGAAATGAGCTCAATGAGGCTGTG 375  
DB 4292 GGGAGGACGAGCTGTGAGGACATCCGATGAATGTGAAATGAGCTCAATGAGGCTGTG 4233  
QY 376 TCCATGACTGTGTGAATATTTCCAGGGAATATGTTGCACTTGTCTTGTAGTGGCTTCAATG 435  
DB 4232 TSCATGACTGTGTCAACATCCCTGGCAATACCGGTGTACCTGCTATGATGATGATGCAAC 4173  
QY 436 TGGCTCATGACGCTCATATTTGCTTGTGAGGAGTGCCTGGAGAAACAATGGCGGCT 495  
DB 4172 TGGCACAATGACGAGACCAACTGTCTGGATGTGACGAGTGTGCCGAGGGCAACGGCGGCT 4113  
QY 496 GCCAGCATACCTGTGTCAACGCTCATCGGAGACTATGATGCTGTGCAAGGAGGGGTTTT 555  
DB 4112 GTACAGAGAGCTGTGTCAACATGATGGGAGCTATGATGCTGCACTGCGGGAAGGCTTCT 4053  
QY 556 TCCTGAGTGACATCAGACACCTGCAATTCACGCTCGGAGAGGGCTGAGCTGCAATGA 615  
DB 4052 TCCTCAGCGAACACCACTACCTGTATCCAGGCGCCAGAGAGGAATGAAATGCAATGA 3993  
QY 616 ATAGGATCAGCGCTGTGAGTCACTGTGCAAGGAGGCGCCCAAGGGGAGGCTGCCCTGTG 675  
DB 3992 ACAAGAACACCGCTGTGCCACATTTCCGAGGAGACACCCAGGGGGGTATTCCTCTGTG 3933  
QY 676 AGTCAGGCGCTGTTTTGAGCTGCCCAAGAACACAGAGAGACTGCTATCTTACCTGTAAAC 735  
DB 3932 AATCCGCTCTGCTTGTGAGCTTACCAAGAACCAACCGGACTGTAAATTCACATGCAACT 3873  
QY 736 ATGGGAACGCTGGGTGCCAGACTCTCTGCAACATACAGCCGATGGCCAGAGTGCAGCT 795  
DB 3872 ATGGTAACGCGGCTGCCAGCACACGTGTGATGACACAGACAGGGTCCCGGCTGCGGCT 3813  
QY 796 GCCATCCACAGTCAAGATGCACACAGATGGGAGGAGCTGCCCTGAGCGAGAGACACTG 855  
DB 3812 GCCATATCAAGTTTGTCTCCATACCGAGGGAAGACATGCAATCGGG-----C 3766  
QY 856 TCCTGAGGTACAGAGAGCAACACACATCAGTGTGGATGGGATAAAGCGGTGAAC 915  
DB 3765 ----GAAAGCGGCTAGAGCAGCACATCCCACT-----C 3735  
QY 916 GCGGCTGCTCATGGAAACGTTGCTGTCTCAACAAATGGAGCTGTGACCGACCTGTAAAG 975  
DB 3734 AAGCGCTTCTTAATGAGACCTGTGCTGTCAACAAACGGGGGCTGTGACAGTAAGTGCCATG 3675  
QY 976 ATACTTCGACAGGTGTCCAGTGCAGTTGTCTGTGATTCACCTCTCCAGTGTGATGGGA 1035  
DB 3674 ATGACGCACTGTGTCTCCACTGCACCTGCCCTGTGGGCTTCACTGTGACCCAGAGAGGA 3615  
QY 1036 AGCATGTAAAGATATTGATGATGAGTGCACAGCCGCAATGGAGGTTGTGATCATTTCTGCA 1695  
DB 3614 AGAGTGCAAAGATATAGATGAGTGGCGCTTAAACAAACGGGGGCTGTGACCATATTGCC 3555  
QY 1096 AAACATCGTGGGAGCTTTTGTACTGGGCTGCAAGAAAGATTTAAATTTAATTAACAGATG 1155  
DB 3554 GCAACACAGTGGGAGCTTCGAAATGCAAGTTGCAAGAAAGCTATAAGCTTCTCATCATG 3495  
QY 1156 AGAAGTCTTCCAGAGATGTGGATGAGTGTCTTGTGATAGAGCTGTGACACAGCTGCA 1215  
DB 3494 AGAGAACTGCCAGGATATAGACGAGTGTCTTGTATGAACTGTGACACACATATGTG 3435  
QY 1216 TCAACACCCCTGGCACATTTGCTTGTGCTTGCATGCAACCGAGGATACCCCTGTATGGCTTCA 1275  
DB 3434 TCAACACACAGAGAGCTTCCAGTGTCTGCTGCCATCGTGGCTACCTGTTGTATGGTATCA 3375  
QY 1276 CCCATGTGGAGACACCAATGAGTGCAGCATCAACAAACGAGGCTGTGACAGAGGCTCTGTG 1335



3374 CCCACTGTGGGATGTGGATGAATGAGCATCAACGGGGAGGTTGCCGCTTTGGCTGCA 3315  
1336 TGAACACAGTGGGAGCTATGATGCGAGTGCACCCCTGG----GTACAGGCTCCACTGGA 1392  
3314 TCAACACTCCTGGCAGCTACCAAGTGTACTGTCCGACAGCGCCAGGGTGGCTGCACTGGA 3255  
1393 ATAAAAAAGACTGTGTGGAAAGTGAAGGGGCTCTCTGCCCAACAAGTGTGTCACCCCCTGTGT 1452  
3254 ATGGCAAGATTGCACAGAGCCACTGAAGTGTACGGGCACTCTGGGCTCTGGAAGCCA 3195  
1453 CCTGCACTGCGGTGAAGTGTGTGGAGAGAGCGGTGCTTCTCAGATGTCACTCTGCA 1512  
3194 TCGCTCAGCTTGCACACCGGTCTGGCAAGGACACCTGTGTCCTGACCTGTCTCCTCAGGG 3135  
1513 TTCACCTCTCTTCAGA-----TGTCAACCACTCAGGACAGTGTACCTTTAGCTAA 1566  
3134 CCGGATTTTGCAGAGGCTGCAGTGTCTTCAATTAACACGGGCTCTCTTCAAGATCA 3075  
1567 ATGAAGGCAAGTGTAGTTTGAA-----AAATGCTGAGCTGTTTCCCGAGGGTCTGCAC 1620  
3074 AGGATGCCAAATGCCGTGTGCACCTCCGAAACAAGGCAACACAGAGGAGGTGCACAA 3015  
1621 CAGCACTACACAGAGAGCACTGATTAAGAGAGCTTCCGCTACGTAAACCTTAGAT 1680  
3014 CCACAGGGCCAGGTGTGTGCCCTCTCTGAATGCCAGGTCACTTCACTCCACCTTAAGT 2955  
1681 CGAGCTCTGGCAAGCAAGTCCACAGAGCCCTGGCGACCAAGCAACCCCTTAGGAAATGT 1740  
2954 GTGACTCTCTCGGAAGGGCAAGGGCCGACGGGCCCGGACCCCTCAGGCAAGAGGTCA 2895  
1741 TTA----TCACTGTGAGTTTGAGCTTTGAAACTAAACAAAGAGAGTGTACAGCTTCTTGTG 1797  
2894 CAAGGCTCACTCGGAACCTGGAGGAGAGGTTCAGAGCCGGAAGAACACAGCCAGCTGTG 2835  
1798 ACTGAGCTGATCTGTAAGAGCAACGAGAGAGGCTCGTAAAGCCATCCGACGCTCA 1857  
2834 GGCTGCCCTGCTCCACAGCGAAATGGAACGGCGGCTGAAAGGATCCCTGAAGATGTCTCA 2775  
1858 GAAAGCCGTCCACAGGAGAGCTTTCACCTCCAGCTCTCAGGATGAACCTCGACGTGG 1917  
2774 GAAAGTCCATCAACAGGACCGCTTCTGCTGGCTGGCAGCCCTTGATATGAGCTGG 2715  
1918 CTAAAAAGCCTCCCAAGACATCTGAACGCCAGGCA-----GAGTCTCTGTGAGTGG 1968  
2714 CCCACAAGCCGGCTGTGTAGCGGGGACGAGAGCGGATGAGTCTGTGTAGCGCCG 2655  
1969 GCCAGGTCTACAGAAACCAATGTGTGAGTGTGAGGCTGGGACCTATTATGTGGAG 2028  
2654 GGCAGCACCGTGTGGACCAAGTGTGTGAGTGTGCGCGAGGAAACGATTACCAAGGCC 2595  
2029 CACGAGAACGCTGCATTTTATGTCCAAATGGAACCTTCCAAATGAGGAGGACAAATGA 2088  
2594 AGACGAGAGTGTGTGCCATGCCAGCGGACCTTCCAGAGAGAGAGAGGCGAGCTCT 2535  
2089 CTGTGAAACCATGCCCAAGACCAAGAAATCTTGGGSCCTTGAGACCCCAAGCTTGA 2148  
2534 CCTGCGACCTTTGCCCTGGGA-----GTGATGCCCAAGCGGCTCTTGGAGCCACA 2484  
2149 ATATGTCTGAATGTGAGAGTGTGTCAACCTGTGTAATATTCTGCAGATGCTTGGAC 2208  
2483 AGCTACACCGTGTGCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2424  
2209 CTGCGAGCTGTGTGCCCTGGGACGCTTCCAGCTTGAAGCTGTGAGCTTCTGCTTCC 2268  
2423 CTTGTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2364  
2269 CTTGTGAGGAGGCTTGGCAACCAACATCAGGAGTCTACTTCTTTCAGGAGTGTGAA 2328  
2363 CTTGTGAGGAGGCTTGGCAACCAACATCAGGAGTCTACTTCTTTCAGGAGTGTGAA 2304  
2329 CAGAGTTCAATGTTTCACTGAGCATTTTACAAACACCACTCAGCATGATTTTCTGTT 2389

2303 CCAAAGTCCAGTCTCCCCAGGCGCACTACTTCAACACCAAGCTCCACCGCTGTATTGGCT 2244  
2389 GCCAGTGGGAAACATACAGCCTGAAATTTGGAAAAAATAATTTGTGTTTCTTGGCCAGGAA 2448  
2243 GTGCAATGGGCTCTCTAAGCCGCACTTCCGTCAGAACTTCTGCGCGGCTGTCCAGGAA 2194  
2449 ATACTAGACTGACTTTGATGGCTCCCAACAACTAACCCAGTGTAAAAAAGAGATGTG 2509  
2183 ACACAAGCAGAGATTTGATGGCTCTACAGTGTGGCCCAATGCAAGATCTGTCAGTGTG 2124  
2509 GAGGGAGCTGAGATTTCACTGGGTACATTGAAATCCCAAACTACCCAGGCAATTTACC 2568  
2123 GTGGGAGTGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2064  
2569 CAGGCAACACCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2628  
2563 CAGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2004  
2529 TGGTCCCTGAGATTTCTCTCCCATAGAGAGCACTGTGGGAGTGTGAGTGTGAGTGTGAGTGTG 2888  
2503 TGGTACCGAGATTTCTCTGCCATCTGAGGATGAGTGTGGGAGCTCTCTCTCATGAGAA 1944  
2589 AAACCTCTTCACTGATTTCTGTGACAACTATGAACCTGCGAGACTAGGAGCGGCA 2748  
1943 AGAATCATCCCATCCCTCCATTTACCACTTATGAGACTTCCAGAGCTTACGAGCGTCCA 1884  
2749 TCSCCTTCACTCCAGTCAAGAGCTGTGGATTCAGTTCAAGTTCAGTTCAGTTCAGTTCAGTTC 2808  
1883 TTSCCTTCACTGCGCTTCCAGGAGCTGTGGATTCAGTTCAGTTCAGTTCAGTTCAGTTCAGT 1824  
2809 GCCTAGAGGTTTCCAGTTCCTCATAGTATGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2868  
1823 GCGCCCTGGCTTCCAGATTTCCCTATGTTACCTATGATGAGGACTATGAGCAGCTGGTAG 1764  
2869 AGACATAGTTCAGATGTGGAGGCTCTATGCTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2928  
1763 AGACATGTGCGAGATGGCGGCTCTATGCTTCTGAAATCCCAAGGAGATTTTAAAGG 1704  
2929 ATAAGAACTTATCAAGGCTCTGTTGATGCTCTGGCCCATCCCAAGAACTATTTCAAGT 2988  
1703 ACAAGAGCTCTCAAGGCTTCTTGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1644  
2989 ACACAGCCAGGAGTCCGAGAGATGTTTCCAGATGTTTCCAGATGTTTCCAGATGTTTCCAGAT 3048  
1643 ACACAG---AGAAACAAGAGGAGTGTGCTGCAAAATCTTCAATCAAGCTCTCCGCTCCA 1587  
3049 AGGTGTCAGGTTTTCAGACCTTACAAAT 3078  
1586 AAGTTCCAGCTTCTGAGGCGCTTACAAAT 1557

## RESULT 7

US-09-839-446-11/C  
; Sequence 11, Application US/29839446  
; Publication No. US2003050232A1  
; GENERAL INFORMATION:  
; APPLICANT: GERLACH, VALERIE L.  
; APPLICANT: ELLERMAN, KAREN  
; APPLICANT: MACDUGALL, JOHN R.  
; APPLICANT: SMITHSON, GLENDA  
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND  
; TITLE OF INVENTION: METHODS OF USING THE SAME  
; FILE REFERENCE: 15966-776  
; CURRENT APPLICATION NUMBER: US/09/839,446  
; CURRENT FILING DATE: 2001-04-19  
; PRIOR APPLICATION NUMBER: 60/198,293  
; PRIOR FILING DATE: 2000-04-19  
; PRIOR APPLICATION NUMBER: 60/198,645  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: 60/210,809  
; PRIOR FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: 60/199,476  
; PRIOR FILING DATE: 2000-04-26



```
; PRIOR APPLICATION NUMBER: 60/200,025
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/224,610
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/200,024
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/199,880
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/218,591
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/271,814
; PRIOR FILING DATE: 2001-02-27
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patencin Ver. 2.1
; SEQ ID NO 11
; LENGTH: 4821
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: POLYX
; OTHER INFORMATION: Z97832.B.0.707
US-09-839-446-11

Query Match      28.6%; Score 1069.2; DB 11; Length 4821;
Best Local Similarity 63.1%; Pred. No. 0;
Matches 1836; Conservative 0; Mismatches 993; Indels 81; Gaps 9;

QY 196 CCGCGGGGCGCAGGAGGATGATGAGTGTGCCCAAGGGGTAGATGACTGCCATGCCG 255
DB 4412 CCCCTCTGCCCTGCCAGATGGATGATGTGTGGAGGGACTGACAACTGCCACATCG 4353
QY 256 AGCCCTGTGTGAGAACACACCCACCTCTCTAAGTGTCTCTGCAAGCCTGCTACCAAG 315
DB 4352 ATGCTATCTGCCAGAACACCCGAGGTCTATCAAGTGTCTGCAAGTCTGCTACACAG 4293
QY 316 GGGAGGAGGAGGAGTGTGAGGACATCATGATGATGTGGAAATGAGCTCAATGAGGCTGTG 375
DB 4292 GGGAGGAGGAGGAGTGTGAGGACATCATGATGATGTGGAAATGAGCTCAATGAGGCTGTG 4233
QY 376 TCCATGACTGTTTGAATATTCAGGCAATATTCGTTGCACTTGTGTTTATGGCTTCAATGT 435
DB 4232 TGCATGACTGTGTCAACATCTCCCTGGCAATTAACGGTGTACCTGTATGATGATTCACACC 4273
QY 436 TGGCTATGACGGTGTATATTTGATGTGAGTGTGAGCAAGTGTCTGTGAGAACAAATGCGGCT 495
DB 4172 TGGCAGATGACGAGACAACTGTCTGATGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4113
QY 496 GCGAGCATACCTGTGTCAACGTGTGAGGAGCTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 555
DB 4112 GTCAGAGAGCTGTGTCAACATGATGGGAGCTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 4053
QY 556 TCCATGAGTGAACATACGACACCTGCAATTCACCGCTCGGAAGAGGCGCTGAGCTGATGA 615
DB 4052 TCCATGAGTGAACATACGACACCTGCAATTCACCGCTCGGAAGAGGCGCTGAGCTGATGA 3993
QY 616 ATAGGATCAGGCTGTATGATCATCTGCAAGAGAGGCGCCCAAGGGGCAAGTGTGCTGTG 675
DB 3992 ACAAGAACACGCTGTGCGCCCATTTTGGCGGAGACACCCCAAGGGGGGTATTGCTGTG 3933
QY 676 AGTGAGGCTGTGTTTGGCTGCGCAAGAACACAGAGAGCTGATCTTGAATCTGTACCT 735
DB 3932 ATGCGCTGTGCTGTGCTTACCAAGAACCAACGGAGCTGTAAATTTGATCAATCAACT 3873
QY 736 ATGGGAACGCTGGTGTGCGGAGCACTCTCTGATGATACAGCGGATGCGCCAGAGTGTGAGT 795
DB 3872 ATGGTAACGCGCTGTGCGGAGCACTGTGATGATGATGATGATGATGATGATGATGATGATG 3813
QY 796 GCATCCACAGTACAGATGACACAGATGGAGGAGCTGCTTGTAGCGAGAGGACATG 855
DB 3812 GCCATATCAAGTTTGTGCTTCCATACCGAGGGAAGACATGATCGGG----- 3766
QY 856 TCTGTGAGGTGACAGAGAGCAACACCATCATGATGATGATGATGATGATGATGATGATGATGAT 915
```

```
DB 3765 ----GAAAGGGGCTAGAGCAGACACATCCCACT-----C 3735
QY 916 GGGCGCTGCTCATGGAAACGTGTGCTGTCAACATGAGGCTGTGACCGCACTGTGAAG 975
DB 3734 AAGCGCTTCTATAGAGACCTGTGCTGTCAACAAACGGGGCTGTGACAGTAAGTGCCATG 3675
QY 976 ATACTTTCAGAGGCTGTGCTGTGAGGCTGTGCTGTGAGGCTGTGCTGTGAGGCTGTG 1035
DB 3674 ATGAGGAGCTGTGCTGTGAGGCTGTGCTGTGAGGCTGTGCTGTGAGGCTGTG 3615
QY 1036 AGCATGTAAAGATATTTGATGAGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1095
DB 3614 AGAGTGTAAAGATATTTGATGAGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 3555
QY 1096 AAAACATCTGTGGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1155
DB 3554 GCAACACAGTGGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 3495
QY 1156 AGAGTCTTCCCAAGATGTGAGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1215
DB 3494 AGAGAACTGCCAGGATATAGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 3435
QY 1216 TCAACACCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1275
DB 3434 TCAACACCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 3375
QY 1276 CCCACTGTGGAGACACCAATGAGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1335
DB 3374 CCCACTGTGGAGACACCAATGAGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 3315
QY 1336 TGAACAGTGGGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1392
DB 3314 TCAACACCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 3255
QY 1393 ATAAAGAGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1452
DB 3254 ATGCAAGATGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 3195
QY 1453 CCCTGTGAGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1512
DB 3194 TGTGAGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 3135
QY 1513 TGTGAGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1566
DB 3134 CCCTGTGAGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 3075
QY 1567 ATGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1620
DB 3074 AGGATGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 3015
QY 1621 CAGCACTACAGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1680
DB 3014 CCAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 2955
QY 1681 GCAGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1740
DB 2954 GTGAGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 2895
QY 1741 TTA---TCACCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1797
DB 2894 CAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 2835
QY 1798 ACCTGAGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1857
DB 2834 GGTGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 2775
QY 1858 GAAAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1917
DB 2774 GAAAGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 2715
QY 1918 CTAAAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 1968
DB 2714 CCCAAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTG 2655
```



QY 416 TTGTTTTCATGGCTTCATGTTGGCTCATGAGGTCATTAATTGTTCTTGATGGAGGAGTG 475  
DB 359 CTGCTATGATGGATTCCACCTTGGACATGACGGACACAACTGCTGGATGTGGACGAGTG 418  
QY 476 CCGGAGAACATGCGGCTGCCAGCATACCTGTGTCAAGCTCATGGGGAGCTATGAGTG 535  
DB 419 TGCCGAGGGCAACGGCGCTGTACAGAGCTGTGTCAACATGATGGGCGACTATGAGTG 478  
QY 536 CTGTGCAAGGAGGGTTTTCCTGAGTGACAAATCAGCACACCTGCATTCACCGCTCGGA 595  
DB 479 CCACTGCCGGAAGGCTTCTTCTCAGCGACAAACCAGCATACCTGTATCCAGCGGCCAGA 538  
QY 596 AGAGGGCTGAGCTGCATGAATAGGATCAAGGCTGTASTACATCTGCAAGAGGGCCCC 655  
DB 539 AGAAGGAATGAATTGCATGAACAAGAACCGGCTGTGCCACATTTTCCGGGAGACACC 598  
QY 656 AAGGGGCAAGCTGCGCTGTGAGTGAGGCTGCTGTTTTCAGCTGGCCAAAGAACCCAGAGAGA 715  
DB 599 CAAGGGGGGATTGCTGTGAATGCCGTCTTGCGCTTGGGCTTACCAAGAACCCAGCGGA 658  
QY 716 CTGCATCTTGACCTGTAAACCATGGGAACGGTGGTGCCAGCACTCCTGTGACGATACAGC 775  
DB 659 CTGTAAATTGACATCAACTATGTGTAAACGGGCTGCCAGCACAGCTGTGTATGACACAGA 718  
QY 776 CGATGGCCAGAGTGCGCTGCCATCCAGATCAAGATGACACAGATGGAGGAGGCTG 835  
DB 719 CGAGGCTCCCGGTGCGCTGCATATCAAGTTTGTGTCTCCATACCGGAGGGAAGACATG 778  
QY 836 CTTGAGCGAGGAGACACTGCTCTGGAGTGACAGAGGACACACCATCAGTGGTGGGA 895  
DB 779 CATCGGG-----GAAAGGGCGGTAGAGGACCATCCCACT-----815  
QY 896 TGGGGAATAACGGGTGAACCGCGCTGCTCATGGAACCGTGTGCTCAACAAATGGAGG 955  
DB 816 -----CAAGCGGTTTCTAATCAGACCTGTGCTGTCAACAACGGGG 856  
QY 956 CTGTGACCGCACCTGTAAGGATATCTGCACAGGTGCTCAGCTGCACTTGTCTGTTGGATT 1015  
DB 857 CTGTGACAGTAAGTCCCATGATGCAAGCACTGCTGTGCCACTGCACCTGCCCTGTGGGCTT 916  
QY 1016 CACTCTCCAGTTGGATGGGAACACATGTAAGATATGATGAGTSCCAGACCGCGCAATGG 1075  
DB 917 CATGCTGACCGACAGACAGACCTGCAAGATATAGATGAGTSCCGCTTAACAACAGG 976  
QY 1076 AGGTTTGATCATTTCTGCAAAAACATCGTGGGCGAGTTTGACTCGGCTGCAGAAAGG 1135  
DB 977 GGGCTGTGACCATATTTGCCGCAACACAGTGGGCGAGCTTGAATGCAGTTGCNAGAAAG 1036  
QY 1136 ATTTAAATTAATTAACAGATGAGAAGTCTTGCCAAAGATGAGTGAATGCTCTTTGGATAG 1195  
DB 1037 CTATAAGCTTCTCATCAATGAGAGAACTGCCAGGATATAGACGAGTGTTCCTTTGATCG 1036  
QY 1196 GACCTGTGACACAGCTGCATCAACACCGCTGGACATTTGCTGTTCACACCGAGG 1255  
DB 1097 AACCTGTGACCCACATATGTGTCAACACAGGAGAGCTTCCAGTGCTCTGCCATCTGTGG 1156  
QY 1256 GTACACCTGTATGCTTCAACCCACTGTGGAGACCAATGAGTGCAGCATCAACAACGG 1315  
DB 1157 CTACCTGTGTATGTATCAACCCACTGTGGGATGTGGATGAATGCAGCATCAACCGGG 1216  
QY 1316 AGGCTGTGAGAGGCTGTGTGTGAACACAGTGGGCGAGCTATGATGCCAGTGCACCTGG 1375  
DB 1217 AGGTTGCCGCTTGGCTGTCATCAACACTCCTGGCAGCTACCACTGTACCTGCCAGCAGG 1276  
QY 1376 ---GTACAAAGCTCCACTGGAATATAAAGACACTGTGTGGAAGTGAAGGGCTCTGCCAC 1432  
DB 1277 CAGGGTCCGCTGCACTGGAAATGGCAAGATTGCCAGAGCCACTGAAGTGTGAGGGCAG 1336  
QY 1433 AAGTGTGTCAACCCGCTGTGTCTCCCTGCATCTGCCGTGAAGAGTGGTGAGAGACGGTGCTT 1492  
DB 1337 TCCCTGGGGCTCGAAAGCCATGCTCAGCTGTCAACCGGTCTGCCAAGAGGACACCTGTGC 1396

QY 1493 CCTCAGATGTCACTTGGCATTCACCTCTCTTTCAGATGTCAACA-----CCATCAGGAC 1546  
DB 1397 CTTGAGCTGTCCCTTCAGGSCCGGATTTTSCAGAGGCTGCAGTCTGTCCATTTAAACA 1456  
QY 1547 AAGTGAACCTTTTAAGCTAAATGAAGCAAGTGTAGTTTGA-----AAATCTGAGCT 1600  
DB 1457 ACGGGCTCTCTCAAGATCAGGATGCCAAATGCCGTTTGCACTCGGAACAAGGCCAA 1516  
QY 1601 GTTTCGGAGGGTCTCGACCAACATTAACAGAGAACAAGCTCAGTCAAAAGAGAGCTT 1660  
DB 1517 AACAGAGAGGCTGGCAGAACACAGGCGCAGGTGGTGGCCCTCTCTGAAATGCCAGGT 1576  
QY 1661 CCGCTAGCTAAACCTTACATGGCAGCTTGGCAGCAAGTCCAGAGAGCCCTGGCGGACC 1722  
DB 1577 CAGCTTCATCCAGCTTAAGTGTACCTCTCTCGGAAGGGCAAGGGCGGCGGCGGAC 1636  
QY 1721 AAGCAACCCCTAAGSAAATGTTTA---TCACTGTGTAGTTTGAGCTTGAACCTAACCAAA 1777  
DB 1637 CCTCCAGGCAAAAGGTCACAGGCTCACCTTGGAACTGGAGGCGAGGTCAGACCGA 1696  
QY 1778 GGAGGTGACAGCTTCTGTGACCTGAGCTGCATCGTAAAGCGAAACCGAGAGGGGTCCG 1837  
DB 1697 AGAAACACAGCCAGCTGTGGGCTGCCCTGCCCTCCGACAGCAATGGAACGGCGGTGAA 1756  
QY 1838 TAAAGCATCCGACGCTCAGAAAGGCGCTCCACAGGGAGCAGTTTTCACCTCCAGCTCTC 1897  
DB 1757 AGGATCCCTGAAGATGCTCAGAAAGTCCATCAACAGGAGCCGCTTCTCTGCGCTGGC 1816  
QY 1898 AGGCATGAACCTCGAGTGCTTAAAAGGCTCCAGAACATCTGAACCGCAGGCA-----1952  
DB 1817 AGGCTTGA-TATGAGCTGGCCCAAGCGGCGCTGTAGCGGGAGCGAGCAGGCC 1876  
QY 1953 -----GAGTCTCTGGAGTGGGCCAGGTCATGCAAGAAACCAATGTGCAGTTGCAAGGC 2008  
DB 1877 GATGAGTCTGTAGGCGCGGCGAGCACCGTGTGTGACCAAGTGTGCAGCTGCCCGCA 1936  
QY 2009 TGGAGCTATATATGGAGCAGCAGACGCTGCATTTATGTCCAATGGAACTTCCA 2068  
DB 1937 GGGAAAGTATACCAAGCGGCGAGCAGGAGTGTGTGCAATGCCACAGCGGCGACCTCCA 1996  
QY 2069 AATAGGAAGGACAAATGACTTGTGAACCATSCCAAGACAGGAAATTTCTGGGGCGCT 2128  
DB 1997 GGAGAGAGAGGCGAGCTCTCTTGGACCTTTGCTGCTGGAG-----TATGCGCCA 2047  
QY 2129 GAAGACCCCAAGAGCTTGAATATGTGAATGTGGAGGTCTGTGTCAACCTGGTGAATA 2188  
DB 2048 CGGCGCTCTTGGAGCCACCAACGTCAACAGCTGTGCAGGTGAGTGCCACCTGGCCAAACA 2107  
QY 2189 TCTGAGATGGCTTGCACCTTGCAGCTCTGTGCGCTGGGCGACGTTCCAGCTGAGC 2248  
DB 2109 CTCTGAGATGGGTTAAGCCCTGTGAGGCAATSCCAAGCTGGCAGCTTACCAAGCTGAAGC 2167  
QY 2249 TGGTCAAGCTTCTGCTTCCCTGTGGAGGAGGCTTGGCCACCAACATCAGGAGGCTAC 2308  
DB 2169 AGAGAGGACCTATGCTTCTCTGCTGGGCGCTCCACCAAGCAATGAAGGGCGCAT 2227  
QY 2309 TCTCTTCAAGAGTGTGAACACAGAGTTCATGTTCACTTGGACATTTCTTCAACACCCAC 2368  
DB 2229 TCTCTTCAAGAGTGTGACACCAAGTCCAGTCTCCCCAGGCGCACTACTACAACACCCAG 2287  
QY 2369 CACTCACCGATGATTCGTTGCCAGTGGGAACATACCAAGCTTGAATTTGGAAAAATAA 2428  
DB 2288 CATCCACCGCTGTATTCTGCTGTGCCATGGGCTCTATCAGCCCGAGCTTCGCTCAGAAGCT 2347  
QY 2429 TCTGTTTCTTGCCAGGAAATACTACGACTGACTTGTGAGTCCACAAACATTAACCCA 2488  
DB 2348 CTGAGGCGCTGTCCAGGAAACACAGCAGACACTTTGATGGCTCTACCGAGTGTGCGCCA 2407  
QY 2489 GTGTAAAAACAGAAAGATGTGGAGGAGGTGGAGATTTTCACTGGGTACATTGAAATCCCC 2548  
DB 2408 ATGCAGAAATCGTCAGTGTGGTGGGAGCTGGGTGAGTTCACTTGGCTATATTGAGTCCCC 2467  
QY 2549 AAATACCCAGGCAATTACCCAGCCAAACACCGAGTGTATGCTGAGCAGCATCAACCCACCCC 2608

```
Db 2468 CAACACCCGGCACTACCCAGCTGCTGGAGTGCATCTGGAACATCAACCCCAACC 2527
Qy 2609 CAAGCGCGCATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2668
Db 2528 CAAGCGCAAGATCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2587
Qy 2669 GCATCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2728
Db 2588 GCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2647
Qy 2729 CCAGACCTACCAAGCGCCCATCGCTTCACTCCAGGTCACAAAGAGCTGCTGCTG 2788
Db 2648 CCAGACCTACCAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2707
Qy 2789 CAAGTCCAAATGAAGGAAACAGCGCTAGAGGCTTCCAGGTCCTCATACGTGACAT 2848
Db 2708 CAAGCAAGCGAGGCAACAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2767
Qy 2849 GCATCTACCAAGGAACTCATGAGACATAGTTCGAGATGCGAGGCTCTATGCTG 2908
Db 2768 GCATCTAGCGAGCTGGTAGAAGACATTTGCGAGATGCGCGGCTCTATGCTGCT 2827
Qy 2909 CCATCAGGAAATACTTAAGGTAAGAACTTATCAAGGCTCTGTTTGTGCTGCTG 2968
Db 2828 CCACAGGAGATTTTAAGGCAAGAGCTCATCAAGGCTCTTCTTGTGCTGCTG 2887
Qy 2969 TCCCGAAGACTTTCAAGTACACAGCCAGAGGTCCTGAGAGATGTTTCAAGATGCT 3028
Db 2888 CCCCAGAACTACTTCAAGTACACAG--AGAAACACAGGAGATGCTGCCAAATGCT 2944
Qy 3029 CATCCGATTGCTACGTTCCAAAGTCTCCAGGTTTTCAGACCTTACAAT 3078
Db 2945 CATCAAGCTGCTCCGCTCCAAAGTTTCAGCTTCTGAGGCTTACAAT 2994
```

## RESULT 9

```
US-09-839-446-13
; Sequence 13, Application US/09839446
; Publication No. US20030050232A1
; GENERAL INFORMATION:
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND
; TITLE OF INVENTION: METHODS OF USING THE SAME
; FILE REFERENCE: 15966-776
; CURRENT APPLICATION NUMBER: US/09/839,446
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: 60/198,293
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/198,645
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/210,809
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/199,476
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,025
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/224,610
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/200,024
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/199,880
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/218,591
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/271,814
; PRIOR FILING DATE: 2001-02-27
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
```

```
; LENGTH: 455C
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: POLYX
; OTHER INFORMATION: 297832_B.1
US-09-839-446-13
```

```
Query Match 26.8%; Score 1067.6; DE 11; Length 4550;
Best Local Similarity 63.1%; Pred. No. 3;
Matches 1849; Conservative 0; Mismatches 999; Indels 82; Gaps 10;
```

```
Qy 176 CCGCGCGGCTCGGGGCGTGGCGGGCGGAGGAGGATGATGATGATGATGATGATGATG 235
Db 120 CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 178
Qy 236 GCTAGATGATGCTGCGATGCGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 295
Db 179 GACTGACAACTGCCACATCGATGCTATCTGCCAGAACACCGCGAGGCTCATACAGT 238
Qy 296 CTGCAAGCTTGGCTACCAAGGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 355
Db 239 CTGCAAGCTTGGCTACCAAGGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 298
Qy 356 TGAGCTCAATGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 415
Db 299 AGAGGATATGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 358
Qy 416 TTGTTTGTAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 475
Db 359 CTGCTATGATGGATTCACCTGCGACATGACGACACACACACACACACACACACAC 418
Qy 476 CTGCGAGAACATGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 535
Db 419 TGCAGGAGGGAACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 478
Qy 536 CTGCTGCAAGGAGGAGGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 595
Db 479 CCACTGCGGGAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 538
Qy 596 AGAGGCTTGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 655
Db 539 AGAAGGAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 598
Qy 656 AAGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 715
Db 599 CAAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 658
Qy 716 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 775
Db 659 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 718
Qy 776 CGATGCGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 835
Db 719 GCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 778
Qy 836 CCTGCGAGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 895
Db 779 CATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 815
Qy 896 TGGGATATAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 955
Db 816 CCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 856
Qy 956 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1015
Db 857 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 916
Qy 1016 CACTCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1075
Db 917 CATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976
Qy 1076 AGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1135
```

Db 977 GGGCTGTGACCAATTTTGGCGCAACACACAGTGGCGCAGCTTCGAATGCAGTTGCAAGAAAGG 1036  
Qy 1136 ATTAAATTTAATACAGATGAAAGCTTTGCCAAGATGTGGATGAGTCTCTTTGGATAG 1195  
Db 1037 CTATAGCTTCTCATATGAGAGAACTGGCCAGGATATAGACGAGTGTCTCTTGATCG 1096  
Qy 1196 GACCTGTGACCAAGCTGATCAACACACCCCTGGCACATTTGGCTGTGTGCTGCAACCGAGG 1255  
Db 1097 AACCTGTGACCAACATATGTGTCAACACACAGAGCTTCCAGTGTCTCTGCCATCGTGG 1156  
Qy 1256 GTACACCTGTATGCTTACCCACTGTGGAGACACCAATGATGCCAGCATCAACACGG 1315  
Db 1157 CTACCTGTGTATGATACCCACTGTGGGATGTGGATGAATGCAGCATCAACCGGGG 1216  
Qy 1316 AGGCTGTGACGAGTGTGTGAAACACAGTGGCGAGCTATGAATGCCAGTGGCACCCCTGG 1375  
Db 1217 AGTTGCCGCTTTGGCTGCATCAACACTCTTGGCAGCTTACCAGTGTACTTGGCCAGAGG 1276  
Qy 1376 ---GTACAAAGCTTCACTGGAATATAAAGAACTGTGTGGAAGTGAAGGGCTCTTGGCCAC 1432  
Db 1277 CCAGGTCGGCTGCATCGAATGGCAAGATTGCCACAGGCCACTGAAGTGTCAAGGGCAG 1336  
Qy 1433 AAGTGTGTACCCGCTGTCTCCCTGCACCTCGGTGAAGTGTGTGGAGAGCGGTGCTT 1492  
Db 1337 TCTGGGGCTCGAAAGCCATGCTCAGTGCACACCGGTCTGGCAAGAGGACACCTGTGC 1396  
Qy 1493 CTTGAGATGTCACTCTGGCATTCACCTCTCTTTCAGATGTCACCA-----CATCAGGAC 1546  
Db 1397 CTTGACCTGTCTCCAGGCCGATTTTGGCAGAGGCTGCAGTGTCTCATTAACA 1456  
Qy 1547 AAGTGTAACTTTAAGCTTAATGAAGGCAAGTGTAGTTGAA-----AATGTGAGCT 1600  
Db 1457 ACGGGCTCTTCAAGATCAAGATGCAAGTGCCTTGTGACCTTGGCAAGCAAGAGGCAA 1516  
Qy 1601 GTTTCGGAGGCTCGGACAGGCTTACAGAGGACACAGTGTGCTGTGCTGTGCTGTGCT 1660  
Db 1517 AACAGAGGAGTGGCAGAACACAGGCGCAGGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1576  
Qy 1661 CGGTACGTAACCTTACATGAGCTCTGGCAAGCAAGTTCAGAGGACCTTGGCCGACC 1720  
Db 1577 CACCTTATCCACCTTAAGTGTGACTCTCTCGAAGGCGCAGGCGCGCCGAGC 1636  
Qy 1721 AAGACCCCTAAGGAAATGTTA---TCACTGTGTGAGTTTGAAGCTTGAACCTTAACAAA 1777  
Db 1637 CCTTCAGGCAAGAGGTCACAAGGCTCACCTGTGAACTTGGAGGCGAGGTCAGAGCGGA 1696  
Qy 1778 GGAGGTGACAGCTTCTGTGACCTGAGCTGCATCGTAAAGCGAACCGAGAGCGGCTCG 1837  
Db 1697 AGAACCCAGGACGCTGTGGCTGCGCTGCTCGACAGCGAATGGAAACGGCGGCTGAA 1756  
Qy 1838 TAAAGCCATCCGACGCTCAGAAAGCCGCTCCACAGGAGAGTTCACCTCCAGCTCTC 1897  
Db 1757 AGGATCCCTGAAGTCTCAGAAAGTCCATCAACAGGACCGCTTCTGCTGCGCTGCG 1816  
Qy 1898 AGGCATGAACCTTGAGCTGGCTTAAAGGCTCCGAGACATCTGAACGCGCAGGA----- 1952  
Db 1817 AGGCTCTGATATGAGCTGGCCCAAGCGCGGCTGTGTAGCGGGAGCGAGCAGAGGCC 1876  
Qy 1953 ---GAGTCTGTGGAGTGGCGCAGGTCATGCAGAAACCAATGTGTGAGTTCGAGGGC 2008  
Db 1877 GATGGAGTCTGTAGGCCCGGCGAGGACCGTGTGGGACCAAGTGTGTGAGTGGCCGCA 1936  
Qy 2009 TGGGACCTATTATGATGGAGCAGAGACGCTGCATTTTATGTCCAAATGGAACTTCCA 2068  
Db 1937 GGGACGTTATACCGCGCCAGAGCGAGCAGTGTGTGCCATGCCAGCGGGCACCTTCCA 1996  
Qy 2069 AATAGGAGGACCAATGACTGTGAACCATGCCCAACACAGGAAATCTGGGGGCTT 2128  
Db 1997 GGAGAGAGAGGCGACCTCTCTCGGACCTTGGCTGGGAG-----TGATGCCCA 2047  
Qy 2129 GAAGACCCAGAGCTTGAATATGTCTGAATGTGGAGTCTGTCAACCTGGTGAATA 2188

Db 2248 CGGCGCTCTTGGAGCCACCAAGCTCACCAGTGTGAGGTGAGTGGCCACCTGGCCACA 2107  
Qy 2189 TTTGTCAGATGGCTTTGCACTTGGCAGCTGTGTGCTGGGACGCTTCCAGCTTGAAGC 2248  
Db 2108 CTCTGTAGATGGGTTCAAGCCTGTGAGCATGCCCAGCTGSCACCTACCACTGAGC 2167  
Qy 2249 TGGTCGAACCTTCTGCTTCCCTGTGGAGAGGCTTGGCACAACAATCAGGAGCTAC 2308  
Db 2168 AGACGGAGCCCTATGCTTCCCTGTGGTGGGCGCTCACCAACAAGCTGAAGGGCCAT 2227  
Qy 2309 TTCTTTTTCAGGACTGTGAACAGAGATTCAATGTTTCACTGTGACATTTCTAGAACAC 2368  
Db 2228 TTCTTTTTCAGGACTGTGACACCAAGTCCAGTGTCTCCAGGCGCATCTACTACACAC 2287  
Qy 2369 CACTCACCCGATGATTCGTTTGGCCAGTGGAAACATACACAGCCTGAATTTGGAAAAAATAA 2428  
Db 2288 CATCCACCGCTGTATTCGCTGTGCCATGGCTCTTATCAGCCCGACTTCCGTGAGACTT 2347  
Qy 2429 TTGTGTTTCTGCCAGGAATACTAGGACTGACTTTTGTATGGCTTCCACAAACATTAACCA 2488  
Db 2348 CTGACGCGCTGTCCAGGAAACACAAAGCACAGACTTTGATGGCTCTACCAAGTGTGGCCCA 2407  
Qy 2489 GTGTAAACACAGAGATGTGGAGGGAGCTGGAGATTTTCACTGGGTACATTTCAATTCGCC 2548  
Db 2408 ATGCAAGAAATGTCAGTGTGGTGGGAGCTGGGTGAGTTTCACTGGCTATTTGAGTCCCC 2467  
Qy 2549 AAATACCCAGGCAATTTAGCCAGCCAAACACCGAGTGTAGTGTGACCATCAACCCACCCCC 2608  
Db 2468 CAATACCCGGGCAACTACCCAGCTGGTGGAGTGCATCTGGAACATCAACCCCGCACC 2527  
Qy 2609 CAAGCGCGGATCTGATGCTGCTCCCTGAGATCTTCTGTCGCCATAGAGAGAGCTGTGG 2668  
Db 2528 CAAGCGCAAGATCTTATGCTGTGATACAGAGATCTTCTGCTGCCATCTGAGGATGAGT 2587  
Qy 2669 GGACTATCTGCTGATGTCGGAACACCTCTTCATCAATTCGTGACAAACATGTAACACCTG 2728  
Db 2588 GGAGCTCTGCTGATGAGAAAGAACTCATCCCATCTCTTACACTTATGAGACCTG 2647  
Qy 2729 CCAGACTACGAAACGCGCCATCGCTTCCCTTCCAGGTCAAAGAGCTGTGGATTCAGTT 2788  
Db 2648 CCAGACTACGAGCGTCCCATTTGCTTCACTGCTTCCAGGCTTCCAGAGCTCTGGATCAACT 2707  
Qy 2789 CAAGTCAATGAAGGAAACAGGCTAGAGGTTCCAGGTTCCAGTCCCATAGCTACATATGATGA 2848  
Db 2708 CAAGCAAGCGAGGCAACAGCGCGCTGCTTCCAGATTCCTATGTTACCTATGATGA 2767  
Qy 2849 GGAATACCGAACTCATTTGAAGACATAGTTCGAGATGGGAGCTCTATGATCTCTGGCCCA 2908  
Db 2768 GGAATACCGAGCTGGTGAAGACATGTTGAGATGGCGGCTCTATGCTCTGGAATA 2827  
Qy 2909 CCATCAGGAAATACTTAAGGATAAGAACTTATCAAGGCTCTTGTATGATCTCTGGCCCA 2968  
Db 2828 CCACCGAGGATTTTAAAGCAAGAGCTCATCAAGGCTTCTTTGAGGTGCTAGCCCA 2887  
Qy 2969 TCCCGAGAACTTTCAAGTACACGCCAGGAGTCCCGAGAGATGTTTCCAGATCGTT 3028  
Db 2888 CCCCCAGAACTTCAAGTACACAG---AGAAACACAAAGGAGATCTGCCAAAATCTTT 2944  
Qy 3029 CATCCGATTCCTACGTTCCAAAGTTCAGGTTTTCAGACCTTACAAAT 3078  
Db 2945 CATCAAGCTGCTCGCTCCAAAGTTTCCAGCTTCTCGAGGCCCTACAAAT 2994

## RESULT 10

US-0-95-142-1  
; Sequence : Application US/10:95142  
; Publication No. US2003036163A  
; GENERAL INFORMATION:  
; APPLICANT: Myriad Genetics, Inc.  
; APPLICANT: Wettstein, Daniel; Albert  
; APPLICANT: Mauck, Kimberly A  
; TITLE OF INVENTION: NOVEL PN826 NUCLEIC ACIDS AND USE THEREOF  
; FILE REFERENCE: 1051.01



```
Db 1894 CATCAAGACCCCTGCGAAGTCCATCGCGCGGAGCAGTCTTATGTCTCAGGTCTCAGGCAC 1893
Qy 1904 GAACCTCGACGTGGGTAAAGACCTCCAGAAATCTGAACCCAGGAGAGTCTCTGG 1963
Db 1854 TGAGTACGAGGTAGCCAGAGCCAGCCAGGCGCTGGAGGGCAGGGG---CATGTGG 1910
Qy 1964 AGTGGCCAGGTCATGCGAAGAAACAAATGTGTCTAGTTCGACGGGTGGGACCTATTATGA 2023
Db 1911 CGCAGGCCAGGTGCTACAGGACAGAAATCGCTGTGGCCCTGGCACCCTCGG 1970
Qy 2024 TGGAGCAGGAGACGCTGATTTATGTCCAAATCGAACTTCCAAATAGGAGGAGCA 2083
Db 1971 TGTGTAGCTGCGCCAGTGTGTGTATGTATGTCAGGAAATACACAGGACATGGAAGGCCA 2030
Qy 2084 AATGACTTGTGAACCATGCCAAGACCCAGGAAATTTCTGGGGCCCTGAAGACCCCGAAGC 2143
Db 2031 GCTCAGTTGCACACCGTCCCGCAGCAGCGCGGTGTGTCTG-----CTGGTGC 2081
Qy 2144 TTGGATATGTCTGAATGTGGAGTCTGTGTCAACTGGTGAATTTCTGCGATGCTT 2203
Db 2082 CGCGAACGTGTGGAATGTGGAGGCGAGTGTCTCCAGGCTTCTTCTCGGCGGATGGCTT 2141
Qy 2204 TGCACCTTCCAGCTCTGTGCCCTGGGCACGTTCCAGCCTGAAGCTGTGCGAACTTCTCG 2263
Db 2142 CAAAGCCTGCCAGGCTGCCCGTGGGCACGTACACAGCTGAGCCGGCGGCGCACCGGTG 2201
Qy 2264 CTTCCCTGTGGAGGAGGCTTGGCCACCAACATCAGGAGACTTCTCTTTCAGGACTG 2323
Db 2202 CTTCCCTGTGGAGGAGGTTTGTCTCACCAACACGAGGACACCACTCTCTTCCAGGACTG 2261
Qy 2324 TGAACCCAGGTTCATGTGTACCTGGACATTTCTACACACACGACTCACCAGTAT 2383
Db 2262 CAGGCTAAAGTGTGCACTGTCTCCCGGCGCCACCACTACACACACGACCCACCGTGCAT 2321
Qy 2384 TGTGTGCCAGTGGGAACATACGAGCTGAAATTTGGAAGAAATTAATTTGTGTCTTGTGCC 2443
Db 2322 CCGTGTCCCGTGGGACCTACCGCCGAGTGTGGCCAGACCACTGCATCACCTGTCTCC 2381
Qy 2444 AGGAATATCTACGCTGACTTGTATGTGTCTCAACAAATTAACCCAGTGTAAACAGAG 2503
Db 2382 GGGCAACACACGACAGACTTCGATGGCTCCACCAACGTCACACACTGCAAAACACGCA 2441
Qy 2504 ATGTGGAGGGAGCTGGGAGATTTCACTGGGTACATTTGATCCCCAACTACCCAGGCA 2563
Db 2442 CTCGCGCGCGAGCTTGTGTGACTACACCGGCTACATCGAGTCCCCCACTACCTGGGCA 2501
Qy 2564 TTACCCAGCCAAACCGAGTGTACGTGGACCATCAACCCACCCCAAGCGCGCATCT 2623
Db 2502 CTACCCAGCCNACGCTGAATGGTCTGGCACATCGCACCTCCCCNAGCGCAGATCT 2561
Qy 2624 GATCGTGTCCCTGAGATCTTCTGCCCATAGAGGACGACTGTGGGACTATCTGGTGA 2683
Db 2562 CATCGTGTCCCTGAGATCTTCTGCCCATCGAGGATGATGCGGCGATGTCTGGTCA 2621
Qy 2684 GCGGAAACCTCTTCATCAATTCGTGACACATATGAACCTGCCAGACCTACGAGC 2743
Db 2622 GAGGAAGTGTCTCTCCACGTCTACCCACCTATGAGCTTCCAGACCTACGAGAG 2681
Qy 2744 CCCCATCGCTTTCACCTCCAGGTCAAGAGGCTGTGGATTTCAGTTCAAGTCCAAATGAAG 2803
Db 2682 GCCATCGCTTTCACCTCCCGCTCCCGCAGCTCTGGATCCAGTTCMAATCCAAATGAAG 2741
Qy 2804 GAACAGGCTAGAGGTTTCAGGTCCCATACGTGACATATGATGAGGACTACAGGAACT 2863
Db 2742 CAACAGCGGCAAGGCTTCCAGTGCCTTATGTACCTACGATGAGGACTACAGCAACT 2801
Qy 2864 CATTGAAGACATAGTTCGAGATGGGAGGCTCTATGCATCTGAGAACCATCAGGAATACT 2923
Db 2802 CATAGAGGACATCGTGGCGGATGGGCGCTGTACGCTCGSAGAACCAACAGGAAATTT 2861
Qy 2924 TAAGGTAAGAACTTATCAAGGCTCTGTTTTCATGTCTGTGGCCCATCCCGAAGCTATT 2983
Db 2862 GAAGACAGAGACTGATCAAGGCCCTCTTCGACGTGCTTGGGCGATCCCGAAGTACTT 2921
```

```
Qy 2984 CAAGTACAGACCCAGAGATCCCGAGAGATGTTTCCAAAGATGTTTCATCCGATTTGCTAGC 3043
Db 2922 CAAGTACAGACCCAGAGATCCAAAGAGATGTTTCCACGCTCTTCATCAAACTGCTGG 2981
Qy 3044 TCCAAAGTGTCCAGGTTTTTTGAGAGCTTACAAATGAC 3081
Db 2982 CTCCAAAGTGTCTCGTTCCTCGCGCCCTACAAATAAC 3019

RESULT 11
US-29-855-824-1
; Sequence 1, Application: US/39855824
; Publication No. US20030166048A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: C0001233
; CURRENT APPLICATION NUMBER: US/09/855,824
; CURRENT FILING DATE: 2001-05-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 1
; LENGTH: 3877
; TYPE: DNA
; ORGANISM: Human
US-29-855-824-1
```

```
Query Match 28.2% Score 1054.8; DB 12; Length 3877;
Best Local Similarity 62.3%; Pred. No. 9,5e-312;
Matches 1923; Conservative 0; Mismatches 962; Indels 207; Gaps 9;

Qy 89 CGCGGGCGGCAACCGTCCCGGGCGGCTGGGGGGGCTGCTCTGCTGCTGCTGCTGCTGCTG 148
Db 48 CCGACGCCACAGCGGGGCGCCCGCATTTAGGCAATGGGCGCGCGGCGCTGGCTGGCACTTGTG 107
Qy 149 GGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 208
Db 108 CGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 167
Qy 209 GGAAGGATGTAGATGATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 268
Db 168 AGTCGAGCTGTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 227
Qy 269 GAACACACCCACCTCTTACAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 328
Db 229 GAACACGCCCAAGTCTTACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 287
Qy 329 GTGTGAGGACATCATGAATGTGGAATG---AGCTCAATGGAGGCTGTGTCCATGACTG 385
Db 288 GTGTGAAGACATTCAGCAGTGTGAGAATGACTACTACAATGGGGGCTGTGCTCAGAGTG 347
Qy 386 TTTCAATATTTCAGGCAATTTATCGTTGCACTGTTTGTGATGGCTTCATGTTGCTCATGA 445
Db 348 CATCAACATCCCGGGAACCTACAGGTGTACCTGCTTTGATGGCTTCATGCTGCAACGA 407
Qy 446 CGGTCAATATTGTCTTGATGTGACAGTGTGCTGGAACAATGGCGGCTGCCAGCATAC 505
Db 408 TGGACACACATCTGCTGATGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGT 467
Qy 506 CTGTGTCAAGCTCATGGGAGCTTATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 565
Db 468 CTGCGTCAATGCCATGGCAGCTACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 527
Qy 566 CAATCAGCACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 625
Db 528 CAACAGCAGCATACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 587
Qy 626 CGGCTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 685
Db 588 TGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 647
```





```
Db 2706 GCGCATGGCTTTCACCTCCGCTCCGCAAGCTCTGGATCCAGTTCAAAATCAATGAAG 2765
Qy 2804 GAACAGCGCTAGAGGCTTCCAGGTCCTCCATACCTGACATATGATGAGGACTACCCAGGAACT 2863
Db 2766 CAACAGCGCAAGGCTTCCAACTGCTTCACTACGATGAGGACTACCCAGGAACT 2825
Qy 2864 CATTAAGACATAGTTCGAGATGGGAGGCTCTATGATCTGAGAACCATCAGGAAATACT 2923
Db 2826 CATAGAGACATCTGCGCGATGGGCGCTGTACGCTCGGAGAACCAACAGGAAATTT 2885
Qy 2924 TAAGATAGAAATATTAAGGCTCTGTTTATCTCTCTGGCCCATCCCGAGACTATTT 2983
Db 2886 GAAAGACAAGAGCTGATCAAGGCCCTCTTCAGCTGTCTGGCGCATCCCGAATCTTT 2945
Qy 2984 CAAGTACACAGCCAGGATCCCGAGAGATGTTTCCAAAGATCGTTCATCCGATTCCTACG 3043
Db 2946 CAAGTACACAGCCAGGATCCCAAGGAGATGTTCCACAGGCTCTTCATCAAACTGCTCG 3005
Qy 3044 TTCCAAAGTGTCCAGTGTTCAGACCTTTTACAAATGAC 3081
Db 3006 CTCCTCAAGTGTCTCGTCTCTCGGCCCTTACAAATAAC 3043

RESULT 12
US-10-195-142-3
; Sequence 3, Application US/10195142
; Publication No. US20030036163A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Wettstein, Daniel Albert
; APPLICANT: Mauck, Kimberly A
; TITLE OF INVENTION: NOVEL PN9826 NUCLEIC ACIDS AND USE THEREOF
; FILE REFERENCE: 1051.01
; CURRENT APPLICATION NUMBER: US/10/195,142
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,323
; PRIOR FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3
; LENGTH: 2967
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-195-142-3

Query Match 28.2%; Score 1052.2; DB 14; Length 2967;
Best Local Similarity 62.7%; P-red No 5e-311;
Matches 1903; Conservative 0; Mismatches 923; Indels 207; Gaps 9;

Qy 153 CTGCTGCTCTGCGGGGCGCTCCGCGGGTCCGGGCGCGTCCCGGGGCGGAGGAG 212
Db 37 CTGCTGGCCCTGGGCACACGCGGGGCGCTGGCGGGGCGAGCGGGTCCCGGGTCACTC 96
Qy 213 GATGTAGATGAGTGTCCCAAGGCTAGATGACTGCATGCCAGCGCCCTGTGTAGAAC 272
Db 97 GAGTGGATGAGTGTCTAGAGGCGCAGATGACTGCCACATCGATGCCATCTGTAGAAC 156
Qy 273 ACACCCACCTCTTACAAAGTGTCTCTCAAGCTGGCTTCCAAAGGCGAAGCGAGGAGTGT 332
Db 157 ACGCCCAAGTCTTACAAATGCTCTCTGCAAGCCAGGCTTCAAGGGGAAGGCAAGGATGT 216
Qy 333 GAGGACATCGATGAATGTGGAATG---AGCTCAATGGAGGCTGTGTCCATGACTGTITG 389
Db 217 GAAGACATTGACGAGTGTGAGATGACTACTCAATGGGGGCTGTGTCCACAGTGCATC 276
Qy 390 AATATTCCAGGCAATATCGTTGCATCTGTTTGTAGTGGCTTCATGTTGGCTCATACGCT 449
Db 277 AACATCCCGGGAACTACAGAGGTACCTGCTTGTATGGCTTCATGTGGCAGACATGGA 336
Qy 450 CATAATTGCTTGTATCTGACGAGTCCCTGGAGAACATGGCGGCTGCCAGCATACCTGT 509
Db 337 CACAATCTGCTGTGATGTGGACGAGTGTACAGGCAATAATGGTGGCTGCCAGGAGATCTGC 396
```

```
Qy 510 GTCAACGTGATGGGAGCTATAGAGTCTGCTGCAAGAGAGGGTGTTCCTGAGTGACAT 569
Db 397 GTCAATGCCATGGGCACTACGAGTGTCTGCTGCAAGAGAGGGTGTTCCTGAGTGACAT 456
Qy 570 CAGCAGACCTGCAATTCACCGCTCGGAAGAGGGCTTGAGCTGCAATGAATAGAGGATCAGGC 629
Db 457 CAGCAGACCTGCAATTCACCGCTCGGAAGAGGGCTTGAGCTGCAATGAATAGAGGATCAGGC 516
Qy 630 TGTAGTCACTATCTCCAAAGGAGGCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 689
Db 517 TGTGCTCAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 576
Qy 690 TTTAGTGGGCGCAAGAACACAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 749
Db 577 TTTAGCTTGGGCGCAAGAACACAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 636
Qy 750 TGGCAGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 809
Db 637 TGGCAGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 686
Qy 810 AAGATGCAACAGATGGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 869
Db 697 GCGCTCCACTCAGACGCTGCGACGTG 722
Qy 870 GAGAGCAACACCATCATCAGTGTGGATGGGATTAACGGGTGAAACGGCGGCTGCTCATG 929
Db 723 -----CATC 726
Qy 930 GMAAGTGTGCTGTCAACAATGGAGGCTGTGACGGCACTGTAAAGATCTTCGACAGGT 989
Db 727 GAGAGCTGGCGAGTCAATAACGGAGGCTGCGACGGGACATGCAAGGACACAGGCACTGGC 786
Qy 990 GTCCACTGCACTTGTCTGTTGGATTCACTCTCCAGTGTGATGGGAGACATGTAAAGAT 1049
Db 787 GTGGATGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 846
Qy 1050 ATTGATGAGTGGCAGACCGCAATGGAGGTGTGATGATTTCTGCAAAACATGCTGGGC 1109
Db 847 ATCAACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 906
Qy 1110 AGTTTGACTGCGGCTGCAAGAAAGGATTTAAATTAATAACAGATGAGAGTCTGCAAA 1169
Db 907 AGCTTCAGTGGGCTGCGGAAAGGCTTACAGCTGCTTCAACGAGGAGGACCTGTCAG 966
Qy 1170 GATGTGATGAGTGTCTTGGATAGGACCTGTGACCAAGCTGATCAACACAGCTGGC 1223
Db 967 GACATCGACGAGTGTCTTGGAGGAGGCTGTGACCAAGCTGATCAACCTGCGGCGGC 1026
Qy 1230 ACATTTGCTTGTCTTGCACCGAGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1289
Db 1027 AGCTTCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1086
Qy 1290 ACCAATGAGTGCAGATCAACAAGGAGGCTGTGACAGGCTGTGCTGCTGCTGCTGCTGCTGCT 1349
Db 1087 GTGACAGGCTGAGCATGAGCAAGGAGGCTGTGACAGGAGGCTGCTGCTGCTGCTGCTGCTGCT 1146
Qy 1350 AGCTATGAATGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1409
Db 1147 AGCTACGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1206
Qy 1410 GA----AGTGAAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1466
Db 1207 GAGACAGGAGTGTCTTCTGCGGCAAGACCTTCCCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1266
Qy 1467 AAGAGTGTGAGGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1526
Db 1267 AAGCAGGCGGCTGTGAGAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1326
Qy 1527 GA----- 1528
Db 1327 GACTCGGAAATAGCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1386
```

QY 1529 -----TGTACCAACATCAGGACAAAGTGTAAACCTTTAAG 1562  
 Db 1397 CAGAAACCAACGGACACAGCTCTGGCCTCGGCCCAAGTGTCTCAGATGCCCCACACACC 1446  
 QY 1563 CTAATAGGAAGCAAGTGTAGTTGAA-----AATGCTGAGCTGTTTCCCGAGGCTGTG 1616  
 Db 1447 CCAATCAACAGAGAGCGCGCTTCAAGATCCAGATGCCAAAGTGCACCTCGCGGCCAC 1566  
 QY 1617 CGACAGCACTACCAAGAGACACAGCT-----CAGTAAAGAGAGCTTCGGTACGTA 1670  
 Db 1507 AGCCAGGACAGAAAGAGAGACCCAGGACCGCTGCTGGACCACTGGCACTGTGACT 1565  
 QY 1671 AACCTTACATGAGCTCTGGCAAGCAAGTCCAGAGAGCCCTGGCGGACCAAGCACCCCT 1730  
 Db 1567 TTGCTGACCTCAAGTGTGACTCTCCAGAGAGGCGCGTGGCGCAAGTCCCACTCC 1626  
 QY 1731 AAGGAATGT-----TTATGACTGTTGAGTTGAACTTAACAAAGAGAGGTGACA 1787  
 Db 1627 AAGGAGGTGTCCACATCACAGCAGAGTTTGAGATCGAGCAAAAGATGGAAGAGCGCTCA 1686  
 QY 1788 GCTTCTGTGACCTGAGCTGCATCGTAAAGCGAACCGAGAGCGGCTCCGTAAAGCCATC 1847  
 Db 1687 GACATATCGAAGCGACTGCTTCGGAAGCGAGCAGACAGAGCTTCGAGCGCCCATC 1746  
 QY 1848 CGACGCTCAGAAAGCGCTCCACAGGAGCAGTTTTCACCTCCAGCTCTCAGGCACTGAAC 1907  
 Db 1747 AAGACCTCGCGCAAGTCCATCGCCCGGAGCAGTGTATGTCAGGCTCTCAGGCACTGAG 1806  
 QY 1908 CTCGAGTGGCTTAAAGGCTCCAGAACATCTGAAGCGGAGAGAGTCTCTGGAGTG 1967  
 Db 1807 TACGAGGTAGCCACAGAGCGACCCAGGCGCTGGAGGGGGCAG---GGGGCATGTGGCGCA 1863  
 QY 1968 GCCCAGGTCATGCAGAAACCAATGTGTCAGTTCGAGGGCTGGGACCTATTATGATGGA 2027  
 Db 1864 GGCCAGTCTACAGCAGCAAAATGCGTTGCTGTGGGCTGGCACCACTTCGTTGTT 1923  
 QY 2028 GCACAGAACCGCTGCTATTTATGTCCAAAATGGAACCTTCCAAATGAGGAAGGACAAATG 2087  
 Db 1924 GAGCTCGGCCAGTGTGTGTCATGTATGCCAGCAACATACACAGGACATGGAAGGCCAGTCC 1983  
 QY 2088 ACTTGTGAACATGCGCAAGACCCAGGAATTTCTGGGCGCTGAAGACCCAGAACCTTGG 2147  
 Db 1984 AGTTGCACACCGTGGCCCGGAGCAGCGGCTTGTCTG-----CTTGTGCGCGC 2034  
 QY 2148 AATATGCTGAATGTGGAGTCTGTGTCAACCTGTGTAATTTCTGCAGATGCTTTGCA 2207  
 Db 2035 AAGTGTGGAATGTGGAGGCAAGTGTCTCCAGGCTTCTTCGGCGAGTGGCTTCAAG 2094  
 QY 2208 CTTTCCAGCTCTGTGCCCTGGGACGTTTCCAGCCTGAAGCTGGTGCAGTCTTCTGCTTC 2267  
 Db 2095 CCCTGCCAGGCTGTGCCCTGGGACGTTACCAAGCTGAGCCCGGCGCACCGCTGCTTC 2154  
 QY 2268 CCCTGGAGGAGGCTTCCGACCAACATCAGGAGCTACTTCTTCAGGACTGTGAA 2327  
 Db 2155 CCCTGTGGAGGGGTTGTCTACCAAAACAGAGCGCACCTCTCTCCAGACTGGCAG 2214  
 QY 2328 ACCAGATTTCAATGTTTCACTGCTGACATTTCTCAACACACACCACTCACCGATGTTCTG 2387  
 Db 2215 GCTAAGTGCATGCTTCCCCCGCCACCACTACACACACCAACCCCGCTGCTGCTCGC 2274  
 QY 2388 TGCCAGTGGGAACATACAGGCTGAATTTGGAATAAATATGTTGTTTTCGCCAGGA 2447  
 Db 2275 TGCCCCGTGGCACCTTACCAGCCCGAGTTGGCCAGAACCACTGCATCACTTCTCCGGGC 2334  
 QY 2448 AATACTACAGTGACTTCTGATGGCTCCACAAACATACCCAGTGTAAAAACAGAGAGTGT 2507  
 Db 2335 AACACAGACAGACTTCGATGGCTCCACCAACGTCACACCTGCAAAACACCAAGCACTGC 2394  
 QY 2508 GGAGGGAGCTGGAGATTTCACTGGGTACATTTGAATCCCAAACTTACCACAGCAATTAC 2567  
 Db 2395 GCGCGGAGCTTGTGACTACACCGGCTACATCGAGTCCCAAACTTACCCTGGGACTAC 2454  
 QY 2568 CCAGCCAAACACCGAGTGTACGTGGACCATCAACCCACCCCGGAGCGGCGCATCTCTGATC 2627

Db 2455 CCAGCCAAACGCTGAATGGCTTGGCACATCGCACCTCCCCAAAGCGCAGGATCCTCATC 2514  
 QY 2628 GTGTCCTCTGAGATCTTCTGSCCATAGAGAGCAGTGTGGGACTATCTGCTGATGGCG 2687  
 Db 2515 GTGTCCTCTGAGATCTTCTGSCCATAGAGAGTGTGGGACTATCTGCTGATGGCG 2574  
 QY 2688 AAAACCTCTTATCCCAATTTCTGTGACAAACATATGAACCTGTGCAGACCTACCAACCGCCC 2747  
 Db 2575 AAGAGTGGCTCTCCACGTCCTATCCACCTATGAGACCTGCCAGACCTACGAGAGGCC 2634  
 QY 2748 ATCGCTTACCTTCCAGGTCCCAAGAGCTGTGGATTCAGTTCAAGTCCCAATGAAGGAAC 2807  
 Db 2635 ATCGCTTACCTTCCAGGTCCCAAGGTGCTATGTCACCTAGTGGAGCTACCAAGCAAC 2694  
 QY 2808 AGCGCTAGAGGCTTCCAGGTCCCAAGGTGCTATGACATATGATGAGGACTACCAAGCAAC 2867  
 Db 2695 AGCGGCAAGGCTTCCAGGTGCTATGTCACCTAGTGGAGCTACCAAGCAAC 2754  
 QY 2868 GAAGACATAGTTCAGATGCGAGGCTCTATGTCATCTGAGAACCATCAGGAATACTTAAG 2927  
 Db 2755 GAGGACATCTGCGCGATGGCGCCTGTACGCTCGAGAGACCAACCAAGATTTGAAA 2834  
 QY 2928 GATAAGAACTTATCAAGGCTCTGTTGATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2987  
 Db 2815 GACAAGAGTGTATCAAGGCTCTTCCAGCTGCTGCGCATCCCAAGAACTACTTCAAG 2874  
 QY 2988 TACAGGCGGAGGCTCCGAGAGATGTTTCCAGATGTTTCCAGATGTTTCCAGATGTTTCCAG 3047  
 Db 2875 TACAGGCGGAGGCTCCAGAGGATGTTTCCAGATGTTTCCAGATGTTTCCAGATGTTTCCAG 2934  
 QY 3048 AAAGTGTCTCGGTTCTGAGACCTTACAAATGA 3120  
 Db 2935 AAAGTGTCTCGGTTCTGAGACCTTACAAATGA 2967

RESULT 13

US-09-898-570-15  
 ; Sequence 15, Application US/09898570  
 ; Patent No. US20020123612A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GERLACH, VALERIE L.  
 ; APPLICANT: ELLERMAN, KAREN  
 ; APPLICANT: MACDOUGALL, JOHN R.  
 ; APPLICANT: SMITHSON, GLENDA  
 ; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND  
 ; TITLE OF INVENTION: METHODS OF USING THE SAME  
 ; FILE REFERENCE: 15966-776C1P  
 ; CURRENT APPLICATION NUMBER: US/09/898,570  
 ; CURRENT FILING DATE: 2001-07-03  
 ; PRIOR APPLICATION NUMBER: 60/198,293  
 ; PRIOR FILING DATE: 2000-04-19  
 ; PRIOR APPLICATION NUMBER: 60/198,645  
 ; PRIOR FILING DATE: 2000-04-20  
 ; PRIOR APPLICATION NUMBER: 60/210,809  
 ; PRIOR FILING DATE: 2000-06-09  
 ; PRIOR APPLICATION NUMBER: 60/199,476  
 ; PRIOR FILING DATE: 2000-04-26  
 ; PRIOR APPLICATION NUMBER: 60/200,225  
 ; PRIOR FILING DATE: 2000-04-26  
 ; PRIOR APPLICATION NUMBER: 60/224,610  
 ; PRIOR FILING DATE: 2000-08-11  
 ; PRIOR APPLICATION NUMBER: 60/200,024  
 ; PRIOR FILING DATE: 2000-04-26  
 ; PRIOR APPLICATION NUMBER: 60/199,880  
 ; PRIOR FILING DATE: 2000-04-26  
 ; PRIOR APPLICATION NUMBER: 60/218,591  
 ; PRIOR FILING DATE: 2000-07-17  
 ; PRIOR APPLICATION NUMBER: 60/271,814  
 ; PRIOR FILING DATE: 2001-02-27  
 ; PRIOR APPLICATION NUMBER: 60/215,855  
 ; PRIOR FILING DATE: 2000-07-03  
 ; PRIOR APPLICATION NUMBER: 60/839,446

```

; PRIOR FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 3177
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: POLYX
; OTHER INFORMATION: CG55098-04
; US-09-998-573-15

```

1120	DB	AACTGTGACCAACATATGTGTCAACACACACCGAAGCTTCCAGTGTCTTGCATCTGGG	1185
1256	QY	GTACACCCCTGTATGGCTTCAACCACTGTGGAGACACCAATAGTGCAGCATCAACAACGG	1315
1186	DB	CTACTGTGTATGTATCATCCACTGTGGGGATGTGGATGAATGCAGCATCAACCGGGG	1245
1316	QY	AGCTGTACAGAGGTCTGTGTGAACACAGTGGGCGAGCTATGAATGCCAGTGCACCCCTGG	1375
1246	DB	AGTTTCCCGCTTTGGCTGTGATCAACACTCTCTGGCAGCTACCAAGTGTACCGTGCACAG	1305
1376	QY	---GTACAGCTCCACTGGAAATAAAAAGACGTGTGGGAGTGAGGGGCTCCTGCCAC	1432
1306	DB	CCAGGGTCGGCTGCACCTGGAAATGGCAAGATGTCACAGAGCCACTGAAGTGTACGGGAG	1365
1433	QY	AAGTGTGTCAACCCGTGTCTCCCTGCACCTCGCGTAAAGTGTGTGGAGGAGACGGGTGCTT	1492
1366	DB	TCTTGGGGCTCGAAAGCCATGCTCAGCTGTGCACCGTCTGCAAGAGGACACCTGTGC	1425
1493	QY	CTCAGATGTCACTCTGGCAATTCACCTCTCTCTCAGA	1528
1426	DB	CCTGACCTGTCTCTCAGGCGCCGATTTTGGCAGAGTCTGAGAATGTGCTTCAAGGTGAG	1485
1529	QY	-----	1528
1486	DB	CTGTGGACCCCGAGCCCGAGGGCTGCTCCAGCCCGAGCTGCGCACATGTGGAAACAGCAC	1545
1529	QY	-----TCTCACCACCATCAGGACAAAGTGAACCTTTAA	1562
1546	DB	CACTCCACCACTGCCATGAGCTGCAGTGTGTCTCCATTAAACACGGGCTCTCTTCAA	1605
1562	QY	GCTAATGAAGSCAAGTGTAGTTGAA-----AAATGCTGAGCTGTTCCTCGAGGGTCT	1615
1606	DB	GATCAAGGATGCCAATGCCGTTTGCACCTCGGAAACAAAGCCAAACAGAGAGGAGCTGG	1665
1616	QY	GCACACAGCACTACACAGAAAGCAAGCTCTAGTAAAGAGAGAGCTTCGGCTACGTAAACCT	1675
1666	DB	CAGAATCACAGGCCACAGTGGTGGTSCCCCTGCTGTGAATGCCAGGTCAACCTTCATCCACT	1725
1676	QY	TACATCGAGCTGTGGCAAGCAAGTCCACGAGGCCCTGCGCACCAAGACCCCTTAAGGA	1735
1726	DB	TAAAGTGTACTCTCTCTGAGAGGCGAAGGCTCAGCGGCCGGACCCCTCCAGGCGAAGA	1785
1736	QY	AATSTTTA---TCACGTGTTGAGTTGAGTTGAACTAACAAAAGAGAGGTGACAGCTTC	1792
1786	DB	GGTCACAAAGGCTCACCTCTGAACTGGAGGCGAGGTCAGAGCGGGAAGAACACACAGCCAG	1845
1793	QY	TTGTGACCTGAGTGTGATCTTAAGCGAACCGAGAGCGGCTCCGTAAAGCCATCCGCAC	1852
1846	DB	CTGTGGGCTSCCTGCTCCGACAGCGAATGGAAACCGCGGCTGAAAGAGTCCCTCAAGAT	1905
1853	QY	GCTCAGAAAGGCGTCCACAGGGAGCAGTTTTCACCTCCAGCTCTCAGGCTAAGACCTCGA	1912
1906	DB	GCTCAGAAAGTCCAATCAACACAGGACCGTTCTCTGCTCGCGCTGGCAGGCTCTGATTATGA	1965
1913	QY	CGTGGGTAAAAAGCCTCCAGAACATCTGAACGCCAGGCA-----GAGTCTCTGG	1963
1966	DB	GCTGGCCCAAGCGGGCTCTGTAGCCGGGAGCGGACAGGCGGATGGAGTCTCTGTAG	2025
1964	QY	AGTGGGCGAGGGTCAATGCGAATAACCAATGTGTCACTTGCAGGGCTTGGACCTTATTATGA	2023

Db 2026 GCCCGGACGACCGTGTCTGGACCAAGTGTGTGAGTCTGCCCGCAGGAAAGTATTACCA 2085  
 Qy 2024 TGAGACACGAGACCGTGCATTTTATGTCCAAATGGACCTTCCAAATGAGGAGGACA 2083  
 Db 2086 CGGCCACAGACGAGACGAGTGTGTCCATGCCACCGGACCTTCCAGAGAGAGAGAGGACA 2145  
 Qy 2084 AATGACTTGTGAACCATGCCACAGACAGGAAATTTCTGGGGCCCTGAAGACCCCAAGAGC 2143  
 Db 2146 GCTCTCTCGACCTTTCGCTGGAG-----TGTATGCCACGGCCCTTTGGAGC 2196  
 Qy 2144 TTGGGAATAAGTCTGAATGTGGAGGTCTGTGTCAACTGTGTGAATATTCGAGATGGCTT 2203  
 Db 2197 CACCAACGTCACACAGTGTGTGACGTCAGTGTGCCACCTGTGCCAACACTCTGTAGATGGGTT 2256  
 Qy 2204 TGCACCTTCCAGCTCTGTGCTGTGGCCTGTGGCAGCTTCCAGCTGAGCTGTGGAACTTCTGT 2263  
 Db 2257 CAAGCCCTGTAGCCATGCCACGTGGCACCCTACCAACCTGAAGCAGGACGACCCCTATG 2316  
 Qy 2264 CTTCCCTGTGGAGGAGGCTTTGCCACCAAAACATCAGGAGGTACTTCTCTTCAGAGCTG 2323  
 Db 2317 CTTCCCTGTGGTGGGGCTTCCACCAAGCATGAGGGCCATTCTCTCCAGAGCTG 2376  
 Qy 2324 TGAACACAGAGTTCAATGTTCACCTGGACATTTCTTACACACACCACTCCACCGATGTAT 2383  
 Db 2377 TGACACCAAAAGTCCAGTGTCTCCACGAGCACTACTACACACCAAGCATCCACCGCTGTAT 2436  
 Qy 2384 TCGTTGCCAGTGGGAACATACAGCCTGAATTTGGNAAATAAATTTGTGTTCTTGCC 2443  
 Db 2437 TCGTGTGCCATGGGCTCTTATGAGCCGACTTCCGTCAAGACTTCTGACGCGCTGTCC 2496  
 Qy 2444 AGGAATACTAGCACTGACTTTTATGGCTTCCAAACATACCCAGTGTGAAGACAGAG 2503  
 Db 2497 AGGAACACACACACAGACTTTGATGGCTTCAACAGTGTGGCCCAATGCAAGATCGTCA 2556  
 Qy 2504 ATGTGAGGGAGCTGGAGATTTCACTGGGTACATTGATTCCTCCAAACTACCCAGGCAA 2563  
 Db 2557 GTGTGTGGGAGCTGGGTGAGTTCACTGGCTATATTGAGTCCCCCACTACCCGGGCAA 2616  
 Qy 2564 TTACCCAGCAACACACCGAGTGTACGTGGACCATCAACCCACCCCAAGCGCGCATCT 2623  
 Db 2617 CTACCCAGCTGTGTGGAGTGATCTGGAACTCAACCCCCCACCACCAAGCGCAATCT 2676  
 Qy 2624 GATCGTGTCCCTGAGATCTTCTGCCCCATAGAGGACGACTGTGGGACTATCTGGTAT 2683  
 Db 2677 TATCGTGTACACAGATCTTCTGTCATCTGAGGATGAGTGTGGGACGTCCTGCTCAT 2736  
 Qy 2684 GCGGAACCTCTTCATCCAATCTGTGACACATATGAACCTGCCAGACCTAGAACG 2743  
 Db 2737 GAGAAGAACTCATCCCCATCTCCATTACCACTTATGAGACTGCGAGACCTAGAGCG 2796  
 Qy 2744 CCCCATCGCTTCACTCCAGGTCAAGAGCTGTGGATTCAAGTCCAAATGAAG 2803  
 Db 2797 TCCATTGCTTCACTGCGCTTCCAGGAAGCTGTGGATCAACTCAAGACAGCGAGC 2856  
 Qy 2804 GAACAGCTAGAGGTTCCAGTCCCATACGTGACATATGATGAGACTACCAAGACT 2863  
 Db 2857 CAACAGCGCCGTGGCTCCAGATTCCTATGTATGATGAGGACTATGAGCAGCT 2916  
 Qy 2864 CATTAAGACATAGTTTCGAGATGGAGGCTTATGATCTGAGAACCATCAGGAAATCT 2923  
 Db 2917 GGTAGAAGACATTTGCGAGATGGCGGCTTATGGCTCTGAAACACCAAGAGATTT 2976  
 Qy 2924 TAAGTAAGAAATTTTCAAGGCTCTGTTGATGTCTGTGGCCATCCCCAGAACTATTT 2983  
 Db 2977 AAAGGACAAAGAGCTCATCAAGGCTTCTTTGAGTGTAGTCCACCCCCCAAGACTACT 3036  
 Qy 2984 CAAGTACACAGCCAGAGTCCGAGAGATGTTTCCAAATCGTTTCAATCCGATTCCTAGC 3043  
 Db 3037 CAAGTACAG---AGAAACACAAAGAGATGCTGCCAAAATCTTTCATCAAGCTCTCCG 3093  
 Qy 3044 TTCCAAAGTGTCCAGGTTTTTGAGACTTACAAAT 3078  
 Db 3094 CTCCAAAGTTTCCAGGTTCTGTAGGCGCTTACAAAT 3128

RESULT 14  
 US-09-839-446-15  
 ; Sequence 15, Application US/09839446  
 ; Publication No. US20010050232A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: GELACH, VALERIE L.  
 ; APPLICANT: ELLERMAN, KAREN  
 ; APPLICANT: MACDOUGALL, JOHN R.  
 ; APPLICANT: SMITHSON, GLENDA  
 ; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND  
 ; TITLE OF INVENTION: METHODS OF USING THE SAME  
 ; FILE REFERENCE: 5965-776  
 ; CURRENT APPLICATION NUMBER: US/09/839,446  
 ; CURRENT FILING DATE: 2001-04-19  
 ; PRIOR APPLICATION NUMBER: 60/198,293  
 ; PRIOR FILING DATE: 2000-04-19  
 ; PRIOR APPLICATION NUMBER: 60/198,645  
 ; PRIOR FILING DATE: 2000-04-20  
 ; PRIOR APPLICATION NUMBER: 60/210,809  
 ; PRIOR FILING DATE: 2000-06-29  
 ; PRIOR APPLICATION NUMBER: 60/199,476  
 ; PRIOR FILING DATE: 2000-04-26  
 ; PRIOR APPLICATION NUMBER: 60/200,025  
 ; PRIOR FILING DATE: 2000-04-26  
 ; PRIOR APPLICATION NUMBER: 60/224,610  
 ; PRIOR FILING DATE: 2000-08-11  
 ; PRIOR APPLICATION NUMBER: 60/200,024  
 ; PRIOR FILING DATE: 2000-04-26  
 ; PRIOR APPLICATION NUMBER: 60/199,980  
 ; PRIOR FILING DATE: 2000-04-26  
 ; PRIOR APPLICATION NUMBER: 60/218,591  
 ; PRIOR FILING DATE: 2000-07-17  
 ; PRIOR APPLICATION NUMBER: 60/271,814  
 ; PRIOR FILING DATE: 2001-02-27  
 ; NUMBER OF SEQ ID NOS: 49  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 15  
 ; LENGTH: 3177  
 ; TYPE: DNA  
 ; ORGANISM: Unknown Organism  
 ; FEATURES:  
 ; OTHER INFORMATION: Description of Unknown Organism: POLYX  
 ; OTHER INFORMATION: CG55098-04  
 US-09-839-446-15

Query Match: 25.7%; Score 959.4; DB 11; Length 3177;  
 Best Local Similarity 60.9%; Pred. No. 1.5e-282;  
 Matches 1847; Conservative 0; Mismatches 1001; Indels 187; Gaps 10;  
 Qy 176 CCGCCGGGTCCGGGCGGTGCGCGGGCGGAGGAGGATGATGATGAGTGTGCCCAAG 235  
 Db 149 CAGCCCGCCGCCCTCAGTACAGCAAGCGGC-GCAGATGTGGATGATGATGTGGAGG 207  
 Qy 236 GCTAGATGATGCTCCATCCGAGCCCTGTGTGAGAACACACCACTCTCAAAAGTCTC 295  
 Db 208 GACTGCAAACTGCCCATGCTCTATCTGCCAGAACACCCCGAGGTATACAAAGTGCAT 267  
 Qy 296 CTGCAAGCTGGTACCAAGGGGAAAGGAGGAGTGTGAGGACATCGATGAATGTGGAAA 355  
 Db 268 CTGCAAGTCTGGGTACACAGGGGACGGCAACACTGCAAGAGCGTGGATGAGTGGAGCG 327  
 Qy 356 TGAGCTCAATGGAGGCGGTGTCCATGACTGTTTGAATATTTCCAGGCAATATCGTTGCAC 415  
 Db 328 AGAGGATATATCCAGTGTGTGTGATGCTGTGTCAACATCCCTGGCAATATCCGCTGTAC 387  
 Qy 416 TTGTTTGTAGTGGTTTCACTGTGGTTCATGACGGTCTCAATTTGTTGATCTGGACAGTG 475  
 Db 388 CTCCTATGATGGATTCACCTGCGACATGACGAGACACAACTGTCTGCA-TGTGACAGAGTG 447  
 Qy 476 CTTGGGAACAAATGCGCGGTGCCAGCATACCTGTGTCAACGTCATGGGAGCTATGAGTG 535

Db 448 TGGCGAGGGCAACGGCGGCTGTGAGCAGAGCTGTGTAACATGATGGGAGCTATGAGTG 507  
Qy 536 CTGCTCAAGAGGGGTTTTCTGTAGTGACATCAGCACACCTGCAITCAGCGTTCGGA 595  
Db 508 CCACTCGCGGAAGGCTTCTTCTCAGCAGACACAGCATACCTGTATCCAGCGGCCAG 567  
Qy 596 AGAGGCGCTGAGCTGATGATGAATGAAGATCAGCGCTGTAGTCACATCTGCAAGAGAGCCCC 655  
Db 568 AGAAGGAATGAATTGATGAACAAGAACCAACGCGCTGTGCCACATTTGCGGAGACACC 627  
Qy 656 AAGGGCAGCGCTGCTGTGAGTGCAGGCGCTGTTTGTAGCTGGCCAAAGAACAGAGAGA 715  
Db 628 CAAGGGGGGTATTGCTCTGTGAATGCGCGTTCTGCTTTGAGCTTACCAAGAACCAACGGGA 597  
Qy 716 CTGCATCTTTGACCTGTAACCATGGGAACGGTGGGTGCCAGCACTCTCTGTGAGCATACAGC 775  
Db 688 CTGTAAATTTGACATGCAACTATGTAAGCGCGCTGCCAGCACACCTGTGTATGACACAGA 747  
Qy 776 CGATGCCCCAGAGTGCAGCTGCCATCGACGTACAGATGCAACAGATGCCAGATGGGAGAGCTG 835  
Db 748 GCAGGGTCCCGGTGGGCTGCCATATCAAGTTTGTGCTCATACCGACGGGAAGACATG 807  
Qy 836 CTTTGAGCGAGAGGACACTGTCTGTGAGGTGACAGAGAGCAACACACATCAGTGTGTGA 895  
Db 808 CATCGGG-----GAAAGCGGCTAGAGCAGACATCCCACT-----844  
Qy 896 TGGGGATAAACGGGTGAACGGCGGCTGCTCATGGAACGTGTGCTGTCAACAATGAGG 955  
Db 845 -----CAAGCGGTTTCTAATGAGACCTGTGCTGTCAACAACGGGG 885  
Qy 956 CTGTGACCGCAGCTGTAAGGATACTTGCAGAGTGTCCACTGCAGTGTGCTTGGATT 1015  
Db 886 CTGTGACGTAAGTGCCATGATCAGCGACTGTGTCCACTGCACCTGGCCCTGTGGCT 945  
Qy 1016 CACTCTCCAGTTGGTGGGAAGACATGTAAGATATTGATGAGTGCAGACCGCAATG 1075  
Db 946 CATGCTGCACCCAGACNGGAAGACGTGCCAAGATATAGATGAGTGGCTTAAACAACGG 1205  
Qy 1076 AGGTTGATCATTTCTGCAANAACATCGTGGGAGTTTGACTGGGGCTGCAAGAAAG 1135  
Db 1006 GGGCTGTGACCATATTGCGGCAACACAGTGGGACGCTTCGATGCAATGCAAGAAAG 1065  
Qy 1136 ATTTAAATTTAACAGATGAGAGTCTTCCAGATGTGGAAGAGTGTCTTTGATG 1195  
Db 1066 CTATAAGCTTCTCATCAATGAGAGGAACCTGCCAGGATATAGACGAGCGTTTCTTTGATG 1125  
Qy 1196 GACCTGTGACACAGTGCATCAACCACTGCGCACAATTGCTTGTGCTTGCACACGAG 1255  
Db 1126 AACCTGTGACCATATGTGTCAACACACAGGAAGCTTCCAGTGTCTTGCCATCTGTG 1185  
Qy 1256 GTACACCTGTATGGCTTCAACCACTGTGGAGACACCAATGAGTGCAGCATCAACAACGG 1315  
Db 1186 CTACCTGTTGATGGTATCACCACTGTGGGATGTGGAAGATGCAAGTCAACCGGG 1245  
Qy 1316 AGGCTGTGACAGGCTGTGTGAACACAGTGGGAGGCTGAATGCGCAGTGCACACCTG 1375  
Db 1246 AGGTTGCGGCTTGGCTGTCATCAACACTCTCTGGCAGCTACCACTGACCTGCCAGCAGG 1305  
Qy 1376 ---GTACAAGCTCACTGGAATAAAGACTGTGTGGAAGTGAAGGGCTCTGCGCCAC 1432  
Db 1306 CAGGCTGGCTGCACTGGGAATGGCAAGATGTGCAGAGCCACTGAAGTGTGAGGCGAG 1365  
Qy 1433 AAGTGTGTACCCCGTGTGCTTCCGCTGCACTGCGGTGAAGAGTGTGGAGGACCGGTGCTT 1492  
Db 1366 TCCTGGGCGCTGGAAGCCATGCTCAGTGCACCGCTCTGGCAGAGGACACCTGTGC 1425  
Qy 1493 CCTCAGATGTACTCTGGCATTCACCTCTCTTCA 1528  
Db 1426 CCTCACCTGTCCCTCCAGGGCCGATTTTGGCCAGAGCTGTGAGAAATGGCTTCACGGTGAG 1485  
Qy 1529 -----1528  
Db 1486 CTGTGGGACCCCAAGCCAGGGCTGTCTCAGCCCGAGCTGGCCACAATGGGAACAGCAC 1545

Qy 1523 -----TGTCAACCAATCAGGACAAAGTGTAACTTTAA 1561  
Db 1546 CAACCTCAACCACTCCATGAGGTGAGTGTCTCCATTAAACAAGGGCTCTCTCAA 1605  
Qy 1562 GCTAAATGAAGGCAAGTGTAGTTTGA-----AAATGCTGAGCTGTTTCCGAGGGTCT 1615  
Db 1636 GATCAAGGATGCCAAATGCCCTTTCACCTCGGAACAAAGGCAAAACAGAGGAGGCTGG 1665  
Qy 1616 GCGACCAAGCACTACAGAGAAACAGCTCAGTAAAGAGAGAGCTTCCCTACGTAAGCT 1675  
Db 1666 CAGAACTACAGGGCAGGTGTGCCCCCTGCTGTGAATGCCAGGTCACTTATCCACT 1725  
Qy 1676 TACATGCAAGCTCTGCGAAGCAAGTCCAGAGAGCCCTTGGCCGCAACAGCAACCTTAAGGA 1735  
Db 1726 TAAGTGTACTCTCTCGGAAGGGCAAGGGCCGACGGCCGCGACCCCTCCAGGCAAGA 1785  
Qy 1736 AATGTTTA---TCACTGTTGAGTTTGAGCTTGAACCTAAACAAAGGAGGTGACAGCTTC 1792  
Db 1786 GGTCAAAAGGCTCACCTTGAACCTGGAGGCAAGGTGAGCCGGAAGAAACAAGCCAG 1845  
Qy 1793 TTGTCACCTGAGCTGCATCGTAAAGCGAAACCGAAGAGGGGTCCGTAAAGGCAATCCGCAC 1852  
Db 1846 CTGCGGCTGCCCTGCCCTCCGACAGCAATGGAACGGGGCTGAAGAGATCCCTGAAGAT 1925  
Qy 1853 GCTCAAAAGGGCGTCCAGAGGAGCAGTTTCACTCCAGCTCTCAGGCATGAACCTGA 1912  
Db 1926 GCTCAGAAAGTCCATCAACCAAGACCGCTTCTGCTGCGCTTGCAGGCTTGATTATGA 1965  
Qy 1913 CTGTCCTAAAGGCTCCCAAGACATCTGAAGCCGACGCA-----GAGTCTCTGTGG 1963  
Db 1966 GCTGCCCAACAGCCCGGCTGTGAGCCGGGAGGAGCAGAGCGATGGAGTCTCTGTAG 2025  
Qy 1964 AGTGGCCAGGGCTCATGCAGAAACCAATGTGTGCTCAGTTTGCAGGGCTGGGACCTATTATGA 2023  
Db 2026 GCCCGGCGAGCACCGTGTGGACCAAGTGTGCTCAGCTGCCCGGAGGAACGTTATACCA 2085  
Qy 2024 TGGAGCAGAGAACGCTGCATTTATGTCCAAATGGAACTTCCAAATGAGGAAGACA 2083  
Db 2086 CGGCCAGAGGAGCAGTGTGTCCATGCCAGCGGCGACCTTCCAGGAGAGAGAAGGCA 2145  
Qy 2084 ATGACTTGTGAACATGCCCAAGACAGGAATTCGGGSCCTGGAAGACCCAGAGG 2143  
Db 2146 SCTCTCTGCCACCTTGGCTTGGAG-----TGATGCCACGGGCTCTTGGAGC 2196  
Qy 2144 TGGAAATGTCTGAATGTGAGGTCTGTCAACCTGGTGAATATTCTGCAGAGTGGCTT 2203  
Db 2197 CACCAAGCTCACCACTGTGCAAGTCACTGCGCAACACTCTGTAGATGGGT 2256  
Qy 2204 TGCACCTTGGCAGCTGTGTGCCCTGGGCACTTCCAGCTTGAAGCTGTGCAACTTCCDTG 2263  
Db 2257 CAAGCCCTGTGAGGATGCCACCTGCCACCTTCCAACTTCCAACTTCCAACTTCCAACT 2316  
Qy 2264 CTTCCTCTGTGGAGAGGCTTGGCCACCAACATCAGGGAGTACTTCTTTCAGAGCTG 2323  
Db 2317 CTTCCTCTGTGGGCGCTTCCACCAACAGCTGGAAGGGGCGCACTTCTTCCAACTG 2376  
Qy 2324 TGAAACCAAGATTCAATGTTCACCTGGACATTTTACAAACACCACTTCCACCTTCCACCT 2383  
Db 2377 TGACCAAAAGTCCAGTGTGCCAGGGGCACTACTACAAACACCACTTCCACCTTCCAT 2436  
Qy 2394 TCGTTGCCAGTGGGAACATACCGCTGAAATTTGGAAAAAATAATTTGTTTCTTGGCC 2443  
Db 2437 TCGCTGTGCATGGGCTCCTATCAGCCGACTTCCGTCAAGAACTTCTCAGCGCGTGTCC 2496  
Qy 2444 AGGAATACTACGACTGACTTTGATGGCTCCACAAACATAACCCAGTGTAAAAACAGAG 2503  
Db 2497 AGGAACCAAGCAGACTTTGATGGCTTACCAAGTGTGGCCCAATCAAGAAATCTCA 2556  
Qy 2504 ATGTGGAGGGAGCTGGAGATTTTCACTGGGTACATTTGAATCCGCCAAATACCGAGCAA 2563  
Db 2557 GTGTGGGAGCTGGGTGAGTTCACTGGCTATATTGAGTCCCACTACCGGGCAA 2616

```

QY 2564 TTACCCAGCCAAACACCGAGTGTAGTGGACCATCAACCCACCCCAAGCGCGGATCCT 2623
Db 2617 CTACCCAGCTGGTGTGGAGTGCATCTGGAAACATCAACCCCAAGCGCGGAAATCCT 2676
QY 2624 GATCGTGGTCCCTGAGATCTTCTGCCCAGTACGAGGACGAGTGTGGGACATCTGTGTAT 2683
Db 2677 TATCGTGGTACAGAGATCTTCTGCCATCTGAGGATGAGTGTGGGACGCTCTGTGTAT 2736
QY 2684 GCGGAAACCTCTTCAZTCAATTTCTGTGACAAATATGAAACCTGCGACACCTTACGAAAG 2743
Db 2737 GAGAAAGAACTCATCCCCATCTCCATTACCACTTATGAGACCTGCGACACCTTACGAGCG 2796
QY 2744 CCCATCGCCTTCACTCCAGTCAAGAGCTGTGGATTCAGTTCAAGTCCAAATGAAAG 2803
Db 2797 TCCATTGCTTCACTGCCGCTCCAGGAAGCTCTGTGATCACTTCAAGCAAGCGAGCG 2856
QY 2804 GAACAGCGCTAGAGGTTCCAGGTCCCATACGTGACATATGATGAGGACATCCAGGAAT 2863
Db 2857 CAACAGCGCGCTGGCTTCCAGATTCCTATGTACCTATGATGAGGACATGAGCAGCT 2916
QY 2864 CATTGAAGACATATGTCAGATGGCAGGCTCTATGATCTGAGAACCATCAGGAATATCT 2923
Db 2917 GGTAGAAGACATTTGTGGAGATGGCGGCTCTATGCTCTGAAAGACACAGGAGATT 2976
QY 2924 TAAGGATAAGAACTTATCAAGGCTGTGTTGATGCTCTGCCCATCCCGACACTATTT 2983
Db 2977 AAGGACAAAGAGCTCATCAAGGCTTCTTTGAGTGTCTAGCCACCCCGAGAACTACT 3036
QY 2984 CAAGTACACAGCCAGAGGTCGCGAGAGATGTTTCCAAGATGTTTCCAGATTCGTTATCGATG 3043
Db 3037 CAAGTACACAG--AGAAACAAAGAGAGTGTGCCAATAATCTTCAATCAAGCTGCTCG 3093
QY 3044 TTCCAAAGTCTCAGGTTTGTGACCTTACAAT 3078
Db 3094 CTCCAAAGTTCCAGCTTCTGAGGCGCTACAAT 3128

RESULT 15
US-09-898-570-9
; Sequence 9, Application US/09898570
; Patent No. US20020123612A1
; GENERAL INFORMATION:
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND
; FILE OF INVENTION: METHODS OF USING THE SAME
; FILE REFERENCE: 15966-776CIP
; CURRENT APPLICATION NUMBER: US/09/898,570
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/198,293
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/198,645
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/210,809
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/199,476
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,025
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/224,610
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/200,024
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/199,880
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/218,591
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/271,814
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/215,855
; PRIOR FILING DATE: 2000-07-03

```

```

; PRIOR APPLICATION NUMBER: 09/839,446
; PRIOR FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 9
; LENGTH: 4440
; TYPE: DNA
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: POLYX
; OTHER INFORMATION: 297812_8.0.7C4
US-09-898-570-9

Query Match 25.23; Score 340.8; DB 10; Length 4440;
Best Local Similarity 6.11; Pred. No. 9.7e-277; Indels 192; Gaps 9;
Matches 1791; Conservative 0; Mismatches 947;

QY 176 CCCCGCGGTGGGGCGGTGCGGGGGCGGCGAGGAGATGTAGATGAGTGTGCCCAAGG 235
Db 120 CACCCCGCGCGCGCGCGAGTACAGCAAGCGCG-GCAAGATGTGATGAGTGTGGAGG 178
QY 236 GCTAGATGACTGCCATGCCGACGCCCTGTGTGACACACACCCACCTCTTACAAAGTCTC 295
Db 179 GACTGACAACTGCCACATCATGCTATCTCCAGAACACCCCGAGGTCTACAAGTGCAT 238
QY 296 CTGCAAGCCTGGTACCAGGGGAGGAGGAGGAGTGTGAGGACATCATGATGAATGTGAAA 355
Db 239 CTGCAAGTCTGGTACACAGGGGAGCGCAACACTGCCAAGAGAGTGGATGAGTGGAGCG 298
QY 356 TGAGTCAATCGAGGCTGTGTCCATGACTGTTTGAATATTCCAGGCAATATTGTTGCAC 415
Db 299 AGAGGATTAATCGAGTGTGTGCAATGACTGTGTCAACATCCCTGGCAATTACCGGTGAC 358
QY 416 TTGTTTGCATGGCTTCAATGTTGGCTCATGACGGTCAATTTGCTTCAATGTGAGGAGTG 475
Db 359 CTGCTATGATGGATTCCACCTGGCACATGACGGACACAACTGTCTGGATGTGACGAGTG 418
QY 476 CCTGGAGAACAAATGGCGCTGCCAGCATACCTGTGTCAACGTCACTGGGAGCTATGAGTG 535
Db 419 TGCCGAGGGCAACGGCGCTGTTCAGCAGAGCTGTGTCAACATGATGGCGAGCTATGAGTG 478
QY 536 CTGCTGCAAGGAGGGGTTTTCTTCTGAGTGACAAATCAGCACACCTGCAATCACCCCTCGGA 595
Db 479 CCACTGCGCGGAGAGCTTCTTCTCAGGACACACGACATACCTGTATCCAGCGCCGAGA 538
QY 596 AGAGGCGCTGAGCTGCAATGAATTAAGATTCACGGTGTAGTCAATCTGCAAGGAGGCGCC 655
Db 539 AGAAGGAATGAATTTGCATGAAACAAAGAACCCAGCGCTGTGCCACATTTGCCGGGAGAC 598
QY 656 AAGGGGAGCGCTCGCTGTGAGTGCAGGCTGGTTTTGAGCTGGCCCAAGAACACAGAGAGA 715
Db 599 CAAGGSGGATATGCTCTGTGAATGCTCTGTGGCTTGGCTTACCAAGAACCAACGGA 658
QY 716 CTGCATCTTGACCTGTAAACCATGGAAACGGTGGGTGCCAGACCTCTCTGCAAGATACAG 775
Db 659 CTGTAAATTTGATGCAATGATGTAACCGGCGCTGCCACGACACGCTGTGATGACAG- 717
QY 776 CGATGGCCCAAGTGCAGCTGCCATCCACATCAAGATGCACACAGATGGAGGAGCTG 835
Db 718 ----- 717
QY 836 CTTGAGCGAGAGGACACTGTCTCTGGAGGTGACAGAGAGCAACACACATCAGTGGTGA 895
Db 718 ----- 717
QY 896 TGGGATAAAGCGGTGAAGCGCGCTGCTCATGGAACGTTGTGCTGTCAACAAATGAGG 955
Db 718 ----- 717
QY 956 CTGTGACGCGACCTGTGAAGATATCTTGCAGAGGTGTCCACTGAGTGTGTCTGTGATTT 1015
Db 747 CTGTGACAGTAAAGTSCATGATGAGGAGGAGTGTGTCTCCACTGCACTTGGCTGTGGCTT 1015

```

1016	QY	CACTCTCCAGTTGGATGGGAAGACATGTAAAGATATTTGATGTCGACACCGCAATGG	1075
807	Db		
807	Db	CATGCTGCACCCAGACAGGAAGACGTGCAAGATATAGATGAGTGCCTTAAACAACCG	866
1076	QY	AGGTTGTGATCATTTCTGCAAAAACATCGTGGGCAGTTTGTACTGCGGCTGCAAGAAAGG	1135
867	Db		
867	Db	GGGCTGTGACCATATTTTGGCGAACACAGTGGGCAGCTTCGAATGCAATGCAATTTGCAAGAAAGG	926
1136	QY	ATTTAAATTATTAAACAGATGAGAAGTCTTGCCAAAGATGTGGATGAGTGCTCTTTGGATAG	1195
927	Db		
927	Db	CTATAAGCTTCTCATCAATGAGAGGAACCTGCCAGGATATAGACAGTGTTCCTTGTATCG	986
1196	QY	GACCTGTGACACAGCTGCATCAACCACTCTGGCAGCATTTGTCTTGTGCTTGGAAACGAGG	1255
987	Db		
987	Db	AACCTGTGACCATATGTGTCAACACACACAGGAAGCTTCCAGTGTCTCTGCCATCGTGG	1046
1256	QY	GTACACCTGTATGGCTTCACCACTGTGGAGACACCAATGAGTGCAGCATCAACAACCG	1315
1047	Db		
1047	Db	CTACCTGTGTATGGTATCACCCACTGTGGGATGTGGATGAATGCAGCATCAACCGGGG	1106
1316	QY	AGGCTGTACAGAGGCTGTGTGTGAACACAGTGGGCAGCTATGAATGCCAGTGCACACCTGG	1375
1107	Db		
1107	Db	AGGTTGCCGCTTTGGCTGCATCAACACTCTCTGGCAGCTACCACTGTACTGTGCCAGAGG	1166
1376	QY	---GTACAAGCTCCACTGGAAATAAAAAGAACTGTGTGGAAGTGAAGGGGTCTCTGCCAC	1432
1167	Db		
1167	Db	CCAGGGTCGGCTGCACCTGGAAATGGCAAGATTTGCACAGAGCCACTGAAGTGTCAAGGCGAG	1226
1433	QY	AAAGTGTACACCGGTGTCCCTGCACTGCGGTAGAGTGGTGGAGGAGACGGGTGCTT	1492
1227	Db		
1227	Db	TCTTGGGGCTCGAAAGCCATGCTCAGCTGCACCGGTCTGGCAAGAGACACCTGTGTC	1286
1493	QY	CCTCAGATGTCACTCTGGCATTCACCTCTCTTCAGA-----TGTCAACCACTACAGGAC	1546
1287	Db		
1287	Db	CCTGACCTGTCCCTCCAGGCCCCGATTTTGGCCAGGCTGCAGTGTCTGCATTAACA	1346
1547	QY	AAAGTGTAACTTTAAGCTTAAATGAAGCAAGTGTAGTTTGA-----AAATGCTGAGCT	1600
1347	Db		
1347	Db	ACGGGCTCTCTCAAGATCAAGGATGCCAAATGCCGTTTGCACCTGCGAAACAAGGCAAA	1406
1601	QY	GTTTCCCGAGGCTCGGCACACACACTACAGAGAGACGACAGCTCAGTAAAGAGAGCTT	1660
1407	Db		
1407	Db	AACAGAGGAGGCTGGCGAACAACACAGGGCAGGTGGTGCCTCTGCTCTGAATGCCAGGT	1466
1661	QY	CCGCTACGTAAACCTTACATGTCAGCTCTGGCAAGCAAGTCCCAAGAGCCCTCTGCCGACC	1720
1467	Db		
1467	Db	CACCTTCATCCACCTTAAGTGTGACTCTCTCGAAGGCGAAGGCCGACGGGCCGGAC	1526
1721	QY	AAGCACCCTTAAGAAATGTTTA---TCACTGTGTAGTGTGAGCTTGAAACTAACCAAAA	1777
1527	Db		
1527	Db	CCCTCAGGCAAGAGGCTCAAGAGGCTCACCCCTGGAACTGGAGSCAGAGTCAAGAGCCGA	1586
1778	QY	GGAGGTGACAGCTTCTGTGACCTGAGCTGCATCGTAAAGCGAACCCAGAGAGGGCTCCG	1837
1587	Db		
1587	Db	AGAAACACAGGCAGCTGTGGGTGCGCTCCGCTCCGACAGCGAATGAACCGGGGCTGAA	1646
1838	QY	TAAAGCATCCGACAGCTCAGAAAGGCGCTCCACAGGAGCAGTTTCACCTCCAGCTCTC	1897
1647	Db		
1647	Db	AGGATCCCTGAAGATGCTCAGAAAGTCCATCAACCAGGACCGCTTCTGTGTGGCTTGGC	1706
1898	QY	AGGCATGAACCTCGACGTGGCTAAAAAGCCTCCACAGAACATCTGAACGCCAGGCA-----	1952
1707	Db		
1707	Db	AGGCCTTGATTATGAGCTGGCCCCACAAGCCGGGCTCGTAGCCGGGAGCGAGCAGAGCC	1766
1953	QY	----GAGTCTGTGGAGTGGGCCAGGCTCATGCAAAAAACCAATGTGTCAAGTTCAGTTCAGGGC	2008
1767	Db		
1767	Db	GATGGAGTCTGTAGGCCCGGCGACACCGTGTGGACCAAGTGTGTCACTGTGCTGCTGCCGCA	1826
2009	QY	TGGACCTATTATGATGAGCACGAGAACGCTGCTATTTATGTCCAAATGGAACCTTCCA	2068
1827	Db		
1827	Db	GGGAACGTATTAACAACGGCAGCGAGAGAGTGTGTGCCATGCCACGGGCGACCTTCCA	1886
2069	QY	AAATGAGGAAGGACAAATGACTTGTGTAACCATGCCCCAAGACCCAGGAATTTCTGGGSCCT	2128

D5	1887	GGAGAGAGAGGCGACGCTCTCTGCGACCTTTGCCCTGG-----GAGTGATGCCCA	1937
QY	2129	GAAGAGCCCGAAGAGCTTGGAAATATGTCTGAATGTGGAGGTCTGTGTCAACCTCTGTGAATA	2188
D5	1938	CGGCGCTCTTGGAGCCACCAACGTCACCGTGTGCAGGTCACTGCGCCACCTCGCCCAACA	1997
QY	2199	TTCTGCGAGATGGCTTTTGCACTTTCGACGCTCTGTGCGCTGGGCAAGTTCCAGCTGAGC	2248
D5	1998	CTCTGTAGATGGGTTCAAGCCCTGTGAGCGATGCCACCGTGGCACTTACCAAGCTGAAGC	2057
QY	2249	TGTTGCAACTTCTGCTTCCCTGTGGAGAGGAGCCCTTGCCACCAACATCAGGAGAGCTAC	2308
D5	2058	AGAGCGGACCTATGCTTCCCTTGGGTGGGGGCTCACCAACGATGAAGGGGCCAT	2117
QY	2309	TTCTTTTCAGGACTGTGAACACAGAGTTCAATGTTTCACTGTGACATTTCTTACAACACCAC	2368
D5	2118	TTCTTTCCAAAGACTGTGACACCAAAATCCAGTGTCTCCCGAGGCGACTACTTACAAACACGAG	2177
QY	2369	CACCTACCGAGTATTCGTTTGCCAGTGGGCAATACCAAGCCCTGAATTTTGGAAAAAATAA	2428
D5	2178	CATCCACCGCTGATTGCTGTGCGATGGGCTCCTATCAGCCGACCTTCGTCAGAACTT	2237
QY	2429	TTGTGTTTCTTGCCCGAGGAAATACTACGACTGACTTTGATGGTTCACAAACATAACCCCA	2488
D5	2238	CTGACGCGCTGTCCAGGAAACAACAGCACAGACTTGTATGGTCTACCAAGTGTGGGCCCA	2297
QY	2489	GTGTAAACACAGAAGATGTGGAGGGAGCTGGAGATTTTCACTGGGTACATTCGAATCCCC	2548
D5	2298	ATGCAAGAATCTGTGAGTGTGGTGGGAGCTGGGTGAGTTCACTGGCTATATTGAGTCCCC	2357
QY	2549	AAACTACCCAGGCAATTACCCAGCCCAACACCGAGTGTACGTGACCATCAACCCACCCCC	2608
D5	2358	CAACTACCCGGCAACTACCCASCTGGTGGAGTGCATCTGGAAATCAACCCCCACC	2417
QY	2609	CAAGCGCGCATCTCTGATCTGTGTCTTGAGATCTTCTGCGCCATAGAGACCACTGTGG	2668
D5	2418	CAAGCGCAAGATCTCTATCTGTGTACCAAGATCTTCTGCGCATCTGAGGATGAGTGTGG	2477
QY	2669	GGACTATCTGGTGATGGGAAAACTCTTCATCCAACTCTGTGACAAACATATGAACCTG	2728
D5	2478	GGAGTCTCTCTCATGAGAAAGAACTCATCCCATCTCTCCATTACCACTTATGASACTG	2537
QY	2729	CCAGACCTACGAACGGCCCCATCGCTTCACTTCCAGTCAAGAAGCTGTGGATTCAGTT	2788
D5	2538	CCAGACCTACGAGCGTCCCATGCGCTTCACTGTCGCGTTCAGGAAGCTCTGGATCAAATT	2597
QY	2789	CAAGTCCAAAGAAAGGAACAGCGCTAGAGGTTCCAGGTCCCATACGTGACATATGATGA	2848
D5	2598	CAAGACAAGCGAGCCCAACAGCCCGCTGGCTTCCAGATTCCCTATGTTACCTATGATGA	2657
QY	2849	GGACTACGAGAACTCATTTGAAGACATAGTTCGAGATGGCAGGCTCTATGCATCTGAGAA	2908
D5	2658	GGACTATGACACGCTGTATGAAGACATTTGTGCGAGATGGCGGCTCTATGCCCTCTGAAA	2717
QY	2909	CCATCAGGAAATACTTAAAGGATTAAGAACTTATCAAGGCTCTGTTTGAATGCTCTGGCCCA	2968
D5	2718	CCACCAGAGATTTTAAAGGACAAGAGCTCATCAAGGCCCTCTCTTGAAGTGTGAGCCCA	2777
QY	2969	TCCCGCAAACTATTTCAAGTACACAGCCCGAGGAGTCCCGAGAGATGTTTCCAAAGATCGTT	3028
D5	2778	CCCCAGAACTACTTCAAGTACACAG---AGAAACACAAGGAGATGCTGCAAAATCTTT	2834
QY	3029	CATCCGATCTGCTACGTTCCAAAGTGTCCAGGTTTTTTTGAGACCTTTACAAAT	3078
D5	2835	CATCAAGCTCTCGGTCTCAAAGTTTCCAGCTTCTCGAGGCCCTACAAAT	2884

Search completed: October 22, 2003, 22:02:39  
Job time : 969 secs

GenCore version 5.1.6  
 Copyright (c) 1993 - 2003 Compugen Ltd.  
 OM nucleic - nucleic search, using sw model  
 Run on: October 22, 2003, 17:27:46 ; Search time 238 Seconds  
 (without alignments)  
 6930.457 Million cell updates/sec

Title: US-09-747-371-1  
 Perfect score: 3737  
 Sequence: 1 ggctcgccgcacaccc.....ctaagcactctggagacat 3737  
 Scoring table: IDENTITY\_NUC  
 Gapop 10.0 , Gapext 1.0  
 Searched: 569978 seqs, 220691566 residues  
 Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000  
 Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Issued Patents NA:  
 1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*  
 2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*  
 3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*  
 4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*  
 5: /cgn2\_6/ptodata/2/ina/PTUS COMB.seq.\*  
 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	257	6.9	291	4	US-09-389-681-246
C 2	257	6.9	291	4	US-09-820-4058-246
C 3	257	6.9	291	4	US-09-339-338-246
C 4	257	6.9	291	4	US-09-433-8269-246
C 5	257	6.9	291	4	US-09-604-287A-246
6	213.4	5.7	241	4	US-09-439-313-412
7	213.4	5.7	241	4	US-09-352-616A-412
8	83	2.2	3373	2	US-08-897-443-2
9	78.6	2.1	1531	2	US-08-833-963C-1
10	78.6	2.1	12018	3	US-08-980-514-2
11	76.6	2.0	1254	4	US-09-312-283C-67
12	75	2.0	1260	3	US-09-188-930-67
13	71.8	1.9	2550	4	US-08-884-072-2
14	71.8	1.9	2550	4	US-09-312-168-2
15	70.2	1.9	2461	1	US-08-282-141-1
16	63.6	1.7	1018	3	US-09-188-930-259
17	63.6	1.7	1018	4	US-09-312-283C-259
18	61.4	1.6	3546	3	US-08-872-757-3
19	61.4	1.6	3546	4	US-09-850-048A-3
20	59.4	1.6	2512	4	US-09-248-757-1
21	58.8	1.6	3284	6	5258288-3
22	58.8	1.6	3290	1	US-07-985-691-1
23	58.8	1.6	3290	1	US-08-436-804-1
24	58.8	1.6	3290	1	US-08-267-387-1
25	58.8	1.6	6811	3	US-08-651-472-67
C 26	58.8	1.6	6811	3	US-08-358-928-67
27	56.4	1.5	2397	6	5258288-2

28	55.9	1.5	2288	4	US-09-620-312D-24
29	54	1.4	5089	6	5177-97-31
C 30	53.6	1.4	220	4	US-09-404-879A-340
31	53.4	1.4	3690	3	US-08-991-408-3
32	53.4	1.4	3690	4	US-09-432-473-3
33	53.4	1.4	3919	2	US-08-866-650-4
34	53.4	1.4	3919	2	US-09-021-287-4
35	53.4	1.4	3919	3	US-09-240-473-4
36	53.4	1.4	5145	3	US-08-931-408-1
37	53.4	1.4	5145	4	US-09-432-473-1
38	52.6	1.4	4661	4	US-09-285-385C-3
39	52.6	1.4	477	2	US-08-866-650-2
40	52.6	1.4	477	2	US-09-021-287-2
41	52.6	1.4	4771	3	US-09-240-473-2
42	52.2	1.4	502	4	US-09-285-385C-1
C 43	52	1.4	2056	3	US-09-334-602-2
44	51.4	1.4	1203	3	US-09-886-010-1
45	50.6	1.4	3759	3	US-08-479-722B-3

## ALIGNMENTS

RESULT 1  
 US-09-389-681-246/c  
 ; Sequence 246, Application US/09389681A  
 ; Patent No. 6518237  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Yuqiu, Jiang  
 ; APPLICANT: Dillon, Davin C.  
 ; APPLICANT: Mitcham, Jennifer L.  
 ; APPLICANT: Xu, Jiangchun  
 ; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND  
 ; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE  
 ; FILE REFERENCE: 210121.470C3  
 ; CURRENT APPLICATION NUMBER: US/09/389,681A  
 ; CURRENT FILING DATE: 1999-09-02  
 ; NUMBER OF SEQ ID NOS: 463  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 246  
 ; LENGTH: 291  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1)-(291)  
 ; OTHER INFORMATION: n=A,T,C or G  
 US-09-389-681-246

Query Match:	6.9%	Score 257;	DB 4;	Length 291;
Best Local Similarity:	97.5%	Pred. No. 2.2e-57;		
Mismatches 268;	Conservative	6;	Indels 1;	Gaps 1;
QY	2506	GGCTGGGACCTATATGATGGAGCAAGAGACCG-TGCATTTTATGTCANATGGAACCT	2064	
DB	275	GGCTGGGACCTATATNATNAGCAGGACGCTTSCATTTTATGTCANATGGAACCT	216	
QY	2065	TCCAAATGAGGAGGACAAATGACCTTGTAACCATGCCAAGCAGGAAATTCGGGG	2124	
DB	215	TCCAAATGAGGAGGACAAATGACCTTGTAACCATGCCAAGCAGGAAATTCGGGG	156	
QY	2-25	CCCTGAAGACCCAGAGCTTGAATATGTCGAATGTGGAGGTCTGTGTCAACCTGGTG	2184	
DB	155	CCCTGAAGACCCAGAGCTTGAATATGTCGAATGTGGAGGTCTGTGTCAACCTGGTG	96	
QY	2185	AATATTCGAGATGGCTTGCACCTTGCACGCTCTGTGCCCTGGGACGTTCCAGCCTG	2244	
DB	95	AATATTCGAGATGGCTTGCACCTTGCACGCTCTGTGCCCTGGGACGTTCCAGCCTG	36	
QY	2245	AAGCTGGTGAACCTTCTGCTTCCCTCTGGGAGGA	2279	
DB	35	AAGCTGGTGAACCTTCTGCTTCCCTCTGGGAGGA	1	



Best Local Similarity 98.7%; Pred. No. 4.5e-46;

Matches 225; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 2350 GACATTTCTACAACACCACCACTACACCGATGATTTCGTTGCCAGTGGGAACATACCAGC 2409

DB 15 GACATTTCTACAACACC-CCACTACCGATGATTTCGTTGCCAGTGGGAACATACCAGC 73

QY 2410 CTGAATTTGGAAAAAATAATTGTGTTCTTCCAGGAAATACACGACTTTTGATG 2469

DB 74 CTGAATTTGGAAAAAATAATTGTGTTCTTCCAGGAAATACACGACTTTTGATG 133

QY 2470 GCTCCACAACATACCCAGTGTAAACACAGAGNATGGAGGGAGCTGGAGATTTC 2529

DB 134 GCTCCACAACATACCCAGTGTAAACACAGAGNATGGAGGGAGCTGGAGATTTC 193

QY 2530 CTGGGTACATTCGAATCCCAACTACCCAGCAATTAACCCAGCCAACA 2577

DB 194 CTGGGTACATTCGAATCCCAACTACCCAGCAATTAACCCAGCCAACA 241

# RESULT 8

US-08-897-443-2

; Sequence 2, Application US/08897443

; Patent No. 5981263

; GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Lal, Preeti

; APPLICANT: Corley, Neil C.

; APPLICANT: Shah, Purvi

; APPLICANT: Kaser, Mathew

; TITLE OF INVENTION: HUMAN XATRILIN-3

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/897,443

; FILING DATE: Filed Herewith

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0348 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3373 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: UTRN0T02

; CLONE: 681719

; US-08-897-443-2

Query Match

2.2%; Score 83; DB 2; Length 3373;

Best Local Similarity 50.7%; Pred. No. 1.6e-11;

Matches 253; Conservative 0; Mismatches 240; Indels 6; Gaps 2;

QY 925 TCATGGAACGTTGCTGTCAACAAATGGAGGTTGTGACCGCACCTGTAAAGGATACTTCA 984

DB 890 TCCAGGATCTGTGCCATGGAGGACCACAACTGTGAGCAGCTCTGTGAATGTGCCG 949

QY 985 CAGGTGTCCACCTGCAGTTGCTCTGTGGATTCACTCTCCAGTTGATGGGAAGACATGTA 1044

DB 950 GCTCCTTTCGTCTGCCAGTGCTACAGTGGCTACGCCCTGGCTGAGGATGGGAAGAGGTGTG 1009

QY 1045 AAGATATTGATGAGTGCAGACCCGCAATGGAGGTTGTGATCATTTCTGCAAAAACATCG 1104

DB 1010 TGGCTGTGAGCTACTGTSCCTCAGAAAACACCGATGTGAACATGAGTGTGTAATGCTG 1069

QY 1105 TGGGTCAGTTTTCAGTGGGCTGCAAGAAAGGATTTAAATTATTAACAGATGAGNAGTCTT 1164

DB 1070 ATGGCTCTCTACTTTTGGCCAGTGCCATGAAGGATTTGTCTTTAAACCCAGATAAAAAACGT 1129

QY 1165 GCACAGATGTGGATGAGTG---CTCTTTGGATAGGACTGTGACCAAGCTGTGACCAAGCTGTGACCAAC 1221

DB 1130 GCACAAAGATAGACTACTGTGCTTCNTCTAATCAGGATGTACAGCAGAGTGTGTTAACA 1199

QY 1222 ACCCTGGCACATTTGCTTGTCTTGCACCGAGGTTACACCTGTAT---GGTTTACACC 1278

DB 1190 CAGATGATTCCTATTCCTGCCACTGCTGAAAGGCTTTACCTGAATCCAGATAAGAAA 1249

QY 1279 ACTGTGGAGACACCAATGAGTGCAGCATCAACACGAGGCTGTACAGAGTCTGTGTA 1338

DB 1250 CCTGCAGAGGATCAACTACTGTGCACTGAACAAACCGGCTGTGAGCATGAGTSCGTCA 1309

QY 1339 ACACAGTGGCAGCTATGATGCCAGTGCACCTGGGTACAAAGCTCCACTGGAAATAAA 1398

DB 1310 ACATGGAGAGAGCTACTACTGCGCTGCCACCGTGGCTACTCTTGACCCCAATGGCA 1368

QY 1399 AAGACTGTGTGGAAGTAA 1417

DB 1370 AAACTGCAGCGAGTGA 1388

# RESULT 9

US-08-933-963C-1

; Sequence 1, Application US/08833963C

; Patent No. 5916769

; GENERAL INFORMATION:

; APPLICANT: Olsen, et al.

; TITLE OF INVENTION: Extracellular/Epidermal Growth Factor

; TITLE OF INVENTION: HCABA58X

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Ave

; CITY: Rockville

; STATE: MD

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/933,963C

; FILING DATE: 11-APR-1997

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WC PCT/US96/05033

; FILING DATE: 10-APR-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Brookes, A. Anders

; REGISTRATION NUMBER: 36,373

; REFERENCE/DOCKET NUMBER: P258

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 301-309-8534

; TELEFAX: 301-309-8439

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

```

;
; LENGTH: 1531 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 125..1453
US-08-833-963C-1

Query Match          2.1%; Score 78.6; DB 2; Length 1531;
Best Local Similarity 54.6%; Pred. No. 1.5e-10;
Matches 206; Conservative 0; Mismatches 159; Indels 12; Gaps 2;

QY 210 GAGGATGTAGATGAGTGTGCCCAAGGGCTAGATGACTGCCATGCCAGCCCTCTGTCTAG 269
DB 488 GTGGATGTGGACGAGTGTGCCAGGCCCTGCACGACTGTGCCCCAGCCAGGACTGCCAT 547
QY 270 AACACACCCAGCTCTTACAAGTGTCTGCAAGCTTCCGCTACCAAGGGGAAAGGAGGCAG 329
DB 548 AACATTGCCCTGGCTCTATCAGTGCACCTGCCCTGATGGTTACCGCAAGATCGGGCCCGAG 607
QY 330 TGTGAGGACATCGATGTGGAATGAGCTCAATGAGGCTGTGCCATGACTGTTTG 389
DB 608 TGTGAGGACATAGACAGTG-----CCGCTACCCCTACTGCCAGCAGCGGTGGTG 658
QY 390 AATATTCCAGCAATATTCGTTGCACTTGTGTTTGTGATGGCTTCATGTTGGCTCATGACGGT 449
DB 659 AACCTGCGCTGGCTCTCCGCTGCCAGTGCAGCGCGGCTTCCAGCTGGGGCTTAACAAC 718
QY 450 CATATTGCTTTGATGTGAGAGTGTCTGCAAGCTTCCGCTGAGAAACAATGGCGGCTGCCAGCATACCTGT 509
DB 719 CGCTCTCTGTGTTGATGTGAACAGTG---TGACATGGGGGCGCCCATGCGAGCAGCGGTGC 775
QY 510 GTCACAGCTCATGGGAGCTATGAGTGTGCTGCAAGGAGGGGTTTTCCTGAGTGACAAAT 569
DB 776 TTCAACTCTATGGGACCTTCTGTGCTGCCACAGGCTATGAGCTGCATCGGGAT 835
QY 570 CAGCACACCTGCATTCA 586
DB 836 GCGCTTCTCTGCAGTGA 852

RESULT 10
US-08-980-514-2
; Sequence 2, Application US/08980514
; Patent No. 6004753
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Guebler, Karl J.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN Si-5-ECMP-LIKE PROT
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/980,514
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
```

```

;
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-C436 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-655-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2018 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTNOT13
; CLONE: 2786449
US-08-980-514-2

Query Match          2.1%; Score 78.6; DB 3; Length 2018;
Best Local Similarity 54.6%; Pred. No. 1.7e-10;
Matches 206; Conservative 0; Mismatches 159; Indels 12; Gaps 2;

QY 210 GAGGATGTAGATGAGTGTGCCCAAGGGCTAGATGACTGCCATGCCAGCCCTCTGTCTAG 269
DB 569 GTGGATGTGGACGAGTGTGCCAGGCCCTGCAGACTGTGCCCCAGCCAGGACTGCCAT 628
QY 270 AACACACCCAGCTCTTACAAGTGTCTTCCAGGCTTCCGCTGAGAAACAATGGGGGAAAGGAGGCAG 329
DB 629 AACTTGGCTGGCTCTTATCAGTGCACCTGCCCTGATGGTTACCGCAAGATCGGGCCCGAG 688
QY 330 TGTGAGGACATCGATGTGGAATGAGCTCAATGAGGCTGTGCCATGACTGTTTG 389
DB 689 TGTGAGGACATAGACAGTG-----CCGCTACCCCTACTGCCAGCAGCCCTCTCGTG 739
QY 390 AATATTCCAGGCAATATTCGTTGCACTTGTGTTTGTGATGGCTTCATGTTGGCTCATGACGGT 449
DB 740 AACCTGCTGGCTCTCTCCGCTGCCAGTGCAGCGGGCTTCAGCTGGGGCTTAACAAC 799
QY 450 CATATTGCTTTGATGTGAGAGTGTCCCTGGAGAAACAATGGGGGCTGCCAGCATACCTGT 509
DB 800 CGCTCTCTGTGTTGATGTGAACAGTG---TGACATGGGGGCGCCCATGCGAGCAGCGCTGC 856
QY 510 GTCACAGCTCATGGGAGCTATGAGTGTGCTGCAAGGAGGGGTTTTCCTGAGTGACAAAT 569
DB 857 TTCAACTCTATGGGACCTTCTGTGCTGCCACAGGCTATGAGCTGCATCGGGAT 916
QY 570 CAGCACACCTGCATTCA 586
DB 517 GCGCTTCTCTGCAGTGA 933

RESULT 11
US-09-312-283C-67
; Sequence 67, Application US/09312283C
; Patent No. 6573095
; GENERAL INFORMATION:
; APPLICANT: Watson, James C.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Marison, James G.
; APPLICANT: Kurbie, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated from Skin Cells
; FILE OF INVENTION: 11005.1011c2
; FILE REFERENCE: 11005.1011c2
; CURRENT APPLICATION NUMBER: US/09/312,283C
; NUMBER OF SEQ ID NOS: 425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 67
; LENGTH: 1254
; TYPE: DNA
; ORGANISM: Mouse
US-09-312-283C-67
```

Query Match 2.0%; Score 76.6; DB 4; Length 1254;  
Best Local Similarity 53.0%; Pred. No. 4.5e-10;  
Matches 188; Conservative 0; Mismatches 164; Indels 3; Gaps 1;  
QY 1068 CGCAATGAGGTTGTGATCATTTCTGCAAAACATCGTGGCAGCTTTTGTGATCGGGCTGC 1127  
DB 180 CGCTATGTTTACTGCCAGCAGCTCTGTGCGAATGTTCTCGATCTATCTCTGTACCTGT 239  
QY 1128 AAGAAAGATTAAATTAATTAACAGATGAGAAGTCTTGCCAAAGTGTGGATGAGTGTCT 1187  
DB 240 AACCTGGCTTACCCCTCAACGATGATGCAAGGTCTTCCCAAGATGTAACGAGTGTGAA 299  
QY 1188 TTGGATAGGACTGTGACCAAGTGTGATCAACACCTCGGCACATTTGCTTGTGCTTGC 1247  
DB 300 ACTGAGAAACCCCTGTGTCAGACCTGCGTCAACACCTATGGTCTTTCATCTGCCGCTGT 359  
QY 1248 AACCGAGGTACACCT---GTATGGCTTCCACCACTGTGGAGACACCAATGAGTGCAGC 1324  
DB 360 GACCCAGATATGAATCGAGGAAGATGGCATTCTACTGCATGATATGATGAGTGCAGC 419  
QY 1305 ATCAACACGAGGCTGTACAGAGTCTGTGTGAACACAGTGGCAGCTATGAATGCCAG 1364  
DB 420 TTCTCCGAGTTCCTGTCAACATGAGTGTGTGAACACGCGGCTCATCTCTCTGCTCA 479  
QY 1365 TGCCACCTGGGTACAACTCCACTGGGATATAAAGAGACTGTGGAAGTGAAGG 1419  
DB 480 TGCCCTCCAGGTACGTCTTGTGGAAGATAACCGAAGCTCCAGGATATCAATG 534  
RESULT 12  
US-09-188-930-67  
; Sequence 67, Application US/09188930A  
; Patent No. 6150502  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Orrust, Rene  
; APPLICANT: Murison, James Greg  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; TITLE OF INVENTION: and Methods For Their Use  
; FILE REFERENCE: 11000.1011c1  
; CURRENT APPLICATION NUMBER: US/09/188,930A  
; CURRENT FILING DATE: 1998-11-09  
; NUMBER OF SEQ ID NOS: 348  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 67  
; LENGTH: 1260  
; TYPE: DNA  
; ORGANISM: Rat  
US-09-188-930-67  
Query Match 2.0%; Score 75; DB 3; Length 1260;  
Best Local Similarity 51.8%; Pred. No. 1.2e-09;  
Matches 184; Conservative 4; Mismatches 164; Indels 3; Gaps 1;  
QY 1068 CGCAATGAGGTTGTGATCATTTCTGCAAAACATCGTGGCAGCTTTTGTGATCGGGCTGC 1127  
DB 186 CGCTATGTTTACTGCCAGCAGCTCTGTGCGAATGTTCTCTGGAATCTATTTCTCTGTACCTGT 245  
QY 1128 AAGAAGAGATTAAATTAATTAACAGATGAGAAGTCTTGCCAAAGTGTGGATGAGTGTCT 1187  
DB 246 AACCTGGCTTACCCCTCAACGATGATGGAAGTCTTGCCAGATGTGAACGAGTGTGAA 305  
QY 1188 TTGGATAGGACTGTGACCAAGTGTGATCAACACCTCGGCACATTTGCTTGTGCTTGC 1247  
DB 306 ACTGAGAAACCCCTGTGTCAGACCTGGGTCACACCTATGGTCTTTCATCTGCCGCTGT 365  
QY 1248 AACCGAGGTACACCT---GTATGGCTTCCACCACTGTGGAGACACCAATGAGTGCAGC 1304  
DB 366 GACCCAGATATGAATCGAGGAAGATGGCATTCATGATGATGGAATGAGTGCAGC 425

QY 1305 ATCAACACGAGGCTGTGATCATTTCTGGAACACAGTGGCAGCTATCAATGCCAG 1364  
DB 426 TTCTCGAGTCTCTGTCAACATGAGTGTGGAACACGCGGCTCATCTTCTGCTCA 485  
QY 1365 TGCCACCTGGTTCACAGCTCCACTGGAATAAAGAGACTGTGTGGAAGTGAAGG 1419  
DB 486 TGCCCTCCAGGCTWCKTCTTGTGGAAGATAACCGAAGTCCAGGATATCAATG 540  
RESULT 13  
US-08-884-072-2  
; Sequence 2, Application US/08884072  
; Patent No. 5872234  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.  
; TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/884,072  
; FILING DATE: Herewith  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0333 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2550 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: CORNNOT3  
; CLONE: 45517  
US-08-884-072-2

Query Match 1.9%; Score 71.8; DB 2; Length 2550;  
Best Local Similarity 52.1%; Pred. No. 1.1e-08;  
Matches 185; Conservative 0; Mismatches 167; Indels 3; Gaps 1;  
QY 1068 CGCAATGAGGTTGTGATCATTTCTGCAAAACATCGTGGCAGCTTTTGTGATCGGGCTGC 1127  
DB 898 CGCTATGTTTACTGCCAGCAGCTCTGTGCGAATGTTCTCTGGATCTCTATTCCTGTACATGC 957  
QY 1128 AAGAAGAGATTAAATTAATTAACAGATGAGAAGTCTTGCCAAAGTGTGGATGAGTGTCT 1187  
DB 958 AACCTGGTGTACCCCTCAATGAGGATGGAAGCTCTTGCCAAAGATGTGAACGAGTGTGCC 1017  
QY 1188 TTGGATAGGACTGTGACCAAGTGTGATCAACACCTCGGCACATTTGCTTGTGCTTGC 1247  
DB 1018 AACGAGAACCCCTGCTGCAAAACCTGCTCAACACCTACGGCTCTTTCATCTGCGGCTGT 1077

QY 1248 AACCGAGGTACACCT---GTATGGCTTACCCACTGTGAGACACCAATGAGTGCAGC 1364  
DB 1078 GACCCAGGATGTAACCTTGAGGAAGATGGCTTATTGCAATGATATGACAGGTGCAGC 1137  
QY 1305 ATCAACACGAGGAGGTGTGAGCAGGTCTGTGTGAACACAGTGGGAGCTATGAATGCCAG 1364  
DB 1138 TTCTCTGAGTTCTCTGCCACATGAGTGTGTGAACACAGTGGGAGCTATGAATGCCAG 1197  
QY 1365 TGCCACCTGGGTACAGCTCCACTGGGAATAAAAGAACTGTGTGAAGTGAAGG 1419  
DB 1198 TGCCCTCCAGGTACATCTCTGCTGATGACACCAAGCTGCAAGACATCAAGC 1252

## RESULT 14

US-09-212-168-2  
; Sequence 2, Application US/09212168  
; Patent No. 6303765  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Guegler, Karl J.  
; TITLE OF INVENTION: HUMAN EXTRACELLULAR MATRIX PROTEINS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Inocyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/212,168  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/884,072  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0333 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2550 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: CORNNOT01  
; CLONE: 45517  
US-09-212-168-2

Query Match 1.9%; Score 71.8; DB 4; Length 2550;  
Best Local Similarity 52.1%; Pred. No. 1.1e-08;  
Matches 185; Conservative 0; Mismatches 167; Indels 3; Gaps 1;  
QY 1068 CGCAATGGAGGTGTGATCATTTCTCAAAAACATCGTGGGAGTTTGAACGCGGTGC 1127  
DB 898 CGCTATGGTTACTGCCAGGAGCTCTGTGCAATGTTCCTGGATCTCTATTCTGTACATGC 957  
QY 1128 AAGAAAGGATTTAAATTTAACAAGATGAGAAGTCTTGCAAGATGTGGATGAGTCTCT 1187  
DB 958 AACCTGGTTTACCCTCAATGAGGATGGAAGTCTTGCAAGATGTGACGAGTGTGCC 1017

QY 1188 TTGGATAGACCTTGACCAAGCTGCATCAACCACTTGGACACATTTGCTGTGCTGC 1247  
DB 1018 AACGAGAACCCCTGTGCTGCAAACTCGTCAACACCTACGGCTCTTTATCTGCGCTGT 1077  
QY 1248 AACCGAGGTACACCT---GTATGGCTTACCCACTGTGAGACACCAATGAGTGCAGC 1304  
DB 1078 GACCCAGGATGTAACCTTGAGGAAGATGGCTTATTGCAATGATATGACAGGTGCAGC 1137  
QY 1305 ATCAACACGAGGAGGTGTGAGCAGGTCTGTGTGAACACAGTGGGAGCTATGAATGCCAG 1364  
DB 1138 TTCTCTGAGTTCTCTGCCACATGAGTGTGTGAACACAGTGGGAGCTATGAATGCCAG 1197  
QY 1365 TGCCACCTGGGTACAGCTCCACTGGGAATAAAAGAACTGTGTGAAGTGAAGG 1419  
DB 1198 TGCCCTCCAGGTACATCTCTGCTGATGACACCAAGCTGCAAGACATCAAGC 1252

## RESULT 15

US-06-282-141-1  
; Sequence 1, Application US/08282141  
; Patent No. 5538861  
; GENERAL INFORMATION:  
; APPLICANT: Schneider, Claudio  
; APPLICANT: Varum, Brian  
; APPLICANT: Avanzi, Giancarlo  
; APPLICANT: Brancolini, Claudio  
; APPLICANT: Manfioletti, Guido  
; TITLE OF INVENTION: Stimulating Factor for the AXI Receptor  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Aegen Inc.  
; STREET: 1840 Denaville Drive  
; CITY: Thousand Oaks  
; STATE: California  
; COUNTRY: United States  
; ZIP: 91320-1799  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.3, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/282,141  
; FILING DATE:  
; CLASSIFICATION: 435  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2461 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
US-08-282-141-1

Query Match 1.9%; Score 70.2; DB 1; Length 2461;  
Best Local Similarity 54.1%; Pred. No. 2.9e-08;  
Matches 166; Conservative 0; Mismatches 138; Indels 3; Gaps 1;  
QY 946 ACAATGGAGCTGTGACCGCACCTGTAAGGATATCTCGACAGGTGCCACTGCAGTTGTC 1005  
DB 622 AGAAGGGGGCTGCTCCAGATCTGCCACACAGCGGGTAGCTTCCACTGTTCTCGCC 681  
QY 1006 CTGTTGATTCACCTCCAGTTGGATGGGAAGACATGTAAGATATTGATGAGTGCASA 1065  
DB 682 ACAGCGCTTCGAGCTCTCTCTGATGGCAGGACCTGCCAAGACATAGACGAGTGGCAG 741  
QY 1066 CCGCAATGAGGTGTGATCATTTCTGCAAAACATCGTGGGACGTTTGAATGCGGCT 1125  
DB 742 ACTCGGAGGCTCGGGGAGCGCGCTGCAAGAACTGCCCGGCTCTACTCTGCTCT 801  
QY 1126 GCAAGAAAGGATTTAAATTTAACAAGATGAGAAGTCTTGCAAGATGTGGATGAGTCT 1185  
DB 932 GTACAGGAGGCTTTGGGTACAGCTCCAGAGAGAGGCTTGGCAGATGTGGACGAGTGTG 859

QY	1186	CTTTGGATAGGACCTGTGACACAGCTGCACTACACACCTGGCACATTTGGCTGTGCTT	1245
Db	860	-CTGCAAGGCGCTGTGAGCAGGCTCTGGGTGAACCTCCGAGGAGCTACACCTGCCACT	918
QY	1246	GCAACCG	1252
Db	919	GTGACGG	925

Search completed: October 22, 2003, 18:45:04  
Job time : 241 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2003, 11:24:43 ; Search time 47 Seconds  
(without alignments)  
2044.096 Million cell updates/sec

Title: US-09-747-371-2  
Perfect score: 5605  
Sequence: 1 MGACNRPRGAWAVLLLLL.....PRSFRLRLRSKVRFLRPYK 999

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 76: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	863.5	15.4	1620	2 T27283	hypothetical prote
2	841.5	15.0	1574	2 T13954	MEGF6 protein - ra
3	727	13.0	2918	2 A54105	fibillin-2 precu
4	724	12.9	2907	2 A57278	fibillin-2 precu
5	691	12.3	2871	2 A55567	fibillin 1 - bow
6	691	12.3	3002	2 A47221	fibillin 1 precu
7	690	12.3	2871	2 A55624	fibillin-1 precu
8	687	12.3	741	2 T46488	hypothetical prote
9	624.5	11.1	1394	2 A35626	transforming grow
10	592	10.6	1712	2 A38261	masking protein pr
11	577.5	10.3	1221	2 A49457	fibulin-2 precu
12	570	10.2	1820	2 A55494	latent transformin
13	567	10.1	685	2 S78040	fibulin, splice fo
14	565	10.1	683	2 C36346	fibulin 1 precu
15	560	10.0	1184	2 A55184	fibulin-2 precu
16	558	10.0	705	2 S34968	fibulin, splice fo
17	553	9.9	601	2 B36346	fibulin 1 precu
18	552.5	9.9	712	2 T42990	fibulin 1, splice
19	540.5	9.6	1251	2 A57293	latent transformin
20	530	9.5	1372	2 T25933	hypothetical prote
21	529.5	9.4	798	2 T2793	hypothetical prote
22	526.5	9.4	2321	2 S78549	notch3 protein - h
23	524.5	9.4	2318	2 S45306	fibulin, splice fo
24	514.5	9.2	689	2 T42760	fibulin-1D precu
25	509.5	9.1	589	2 T43210	notch protein - fr
26	509.5	9.1	2703	1 A24420	hypothetical prote
27	507.5	9.1	3507	2 T34513	Xotch protein - Af
28	502.5	9.0	2524	2 A35844	notch protein homo
29	497	8.9	2555	2 A40043	

30	494.5	8.8	2531	2 T31070	notch homolog - se
31	493	8.8	2533	2 A49175	Notch B protein -
32	489	8.7	2531	2 S49198	notch protein homo
33	489	8.7	2531	2 A46019	notch-1 protein -
34	487.5	8.7	1964	2 T09059	notch4 - mouse
35	487	8.7	2352	2 T30201	Notch homolog prot
36	484	8.6	2437	2 S42612	transmembrane prot
37	479	8.5	2471	2 A49129	cell-fate determin
38	466	8.3	1064	2 A40136	fibropellin Ia - s
39	443.5	7.9	1429	2 S06434	homeotic protein 1
40	441.5	7.9	1274	2 T42017	cysteine rich prot
41	438	7.8	1217	1 EGWSMG	epidermal growth f
42	431.5	7.7	1133	1 EGRT	epidermal growth f
43	423	7.5	1827	2 T34288	hypothetical prote
44	422	7.5	3623	2 T38618	intrinsic factor-B
45	421	7.5	1207	1 EGHC	epidermal growth f

ALIGNMENTS

RESULT 1

T27283

hypothetical protein V64G10A.f - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T27283

R:Ainscough, R.

submitted to the EMBL Data Library, September 1999

A:Reference number: 220336

A:Accession: T27283

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1620 <MIL>

A:Cross-references: EMBL:AL110498; NID:e:542303; PIDN:CAB54471.1; CESP:Y64310A.f

A:Experimental source: clone Y64G10A

C:Genetics:

A:Gene: CESP:Y64G10A.f

A:Introns: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 466/1; 549/1; 559/1; 601/1; 625/1

Query Match 15.4%; Score 863.5; DB 2; Length 1620;

Best Local Similarity 24.9%; Pred. No. 1.5e-44;

Matches 280; Conservative 100; Mismatches 372; Indels 371; Gaps 48;

```

Db      519 AQLCKNRKSGRCQCFAGVILAHDEKSCVAASDADISNIEDIYSKVPGLDSIDEVSS 578
QY      449 ---LPTSVSPR-----VSLHCKKGGGCGCFLRCHSGIH::SSDVT 486
Db      579 IESYPADSPRLVFGRRRRHVACVNFQGTLSLELFSSEVRTPDSEKCPNGFFGS--- 634
QY      487 IRTSVTFKLNKCKSLKNAELF-----PEG----- 511
Db      635 CQ::SCSDCOQNGKCKMRGSGLLSKDCPSGYGKECEQ::CRNGYWGVDCAKCKCKL::CDP 694
QY      512 ---LRPALPEKSS--VKESF-----RYVNLTCSSGKQVPGAGRPSTPEKXFITVEF 559
Db      695 STGSCRECEPEKSDGCPDGPYGYGQCNLKRMDCPNRCDB-VFYCYCTPDGLY--- 748
QY      560 ELSTNQKEVTASCDJSCIVKRTKRLRAIRLTK-----AVHEQ 600
Db      749 -----GQCEKPCPHFTFGKNCRFPCKARENSEGCDEITGKCRCKPGYYGHCKR 799
QY      601 F-----HLQSGNLDVAKKPR-----TSBEQA 624
Db      800 MCSPLGLFAGAGAKKSCAPAGIRCDPVTGDTTKKCPAGYCGNLCDQPCPAGYGYDCEQXC 859
QY      625 ESCGVQGHAEHQ-----VSCRAGTY 646
Db      860 SCADVASPHKSKVCHRVGTCTCLPGKTGLCDQGLIFVETIEFDIAFSINVIACPNY 919
QY      647 YDGAERBCILCPNGTQNE-BGOMTCEP-----CP--RPG----- 678
Db      920 GPNCATHC-SCVNGAKCDESCHCTPGFYGATCSEVCPTGRFGIDCMQLCKQNGAIC 978
QY      679 ---NSGALKTPKAWNMSECGGLCOPGEYSADGAPOLCALGTF-OPEAGRTSCP----- 730
Db      979 DTSNGSCCAPGSGKXKDKACAPGTGKDCSKKD-CADGHCDPDSGECICPPKKGH 1037
QY      731 -----CGGSL-----ATKHOGATSFQ---DQETRV-----QCSPGHF----- 759
Db      1038 KCDTCDSLGAGCKGICSCQNGATCDSVTSGCECRPGMKKCDRCPDGRFEGCNA 1097
QY      760 -----YNTTHRC-----IRCPVGYOPEGKNNC-VSCPNTTTFDGSIN 800
Db      1098 ICDTTNTDTSYNPVAVCDHVTGECRCPAGWTGPD-----CQTSCP-LGRHGEGRH 1150
QY      801 ITQCKNRRCGGLDPTGVIESPVPNVPANTEC---TWTIN 840
Db      1151 SCQCSN---GASCDRTGDCDPSGFMGKNCSECEPEGLWSN 1190

RESULT 2
T13954
MEGF6 protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T13954
R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
Genomics 51, 27-34, 1998
A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs
A:Reference number: Z14126; MUID:98360089; PMID:9693030
A:Accession: T13954
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1574 <NAX>
A:Cross-references: EMBL:AB011532; NID:G3449293; PIDN:BAA32462.1; PID:G3449294
A:Experimental source: strain Prague-Dawley; brain
C:Genetics:
A:Gene: MEGF6

Query Match 15.0%; Score 841.5; DB 2; Length 1574;
Best Local Similarity 26.3%; Pred. No. 3e-43;
Matches 245; Conservative 74; Mismatches 289; Indels 323; Gaps 40;

QY      34 .PGRGRAAGPO---EDVDECAQGLDDCHADALQN::PITSYKSCCKEY--QSEGCCEID 88

```

```

Db      109 PCHSKPFGEGCLSDVDEKASANGGC---EGPCNTVGGFYCECPEGYO::QSDGKTCQD 166
QY      89 ECGNELNGCVHDCNLIPONVACTCFDGFYLAHDGHNCLDVDELENNGGQOHTCNV-M 147
Db      167 EC-RAHNGGCGHRCVNTPGSYLCECKPGRJLTDGRTCLAISSCTLGNGGCGHQCVQ 225
QY      148 GSVECCCKEGFE::SDNQHTCIIHRSBGLSQMKDHGCSHICKEAPRGSSVA::ECRPGFELA 207
Db      226 TCRCCRCRQYQLQEDGRRCVRRS---PCAEGNGGCMHICOEL::RGLAHGCGHPGYLA 280
QY      208 XNGRDC--ILTCNNGNGGCGHSCDQDADPECSCHPQVXVHTDGRSCLERDQTVLEVTES 265
Db      281 ADPKTCEDVDEKALG::AQCACPLNTQSFCKVCVAGVELGADGQCRIE---MEIYNS 337
QY      266 NTSVVGDKRVKRLNLETCAVNGGCDRTCKQTS::GVHSCSPVGTJQLDGKTKCKID 325
Db      338 -----CBAGNGGCSHGSHSTSTPLCTCPRGYE::DEQKTCIDID 377
QY      326 ECQTRNGGCDHFCKNIIVSGDCGCKKFKLJLTDKSCQDQDDECSLDR-TCDSHCINHPGT 384
Db      378 DC-ANSPCCQACANTPGGYECSCFAGYRLNTDGGCSDVDEKASGHGCGCHSCNLAGS 436
QY      385 FACACNRGYT----- 400
Db      437 PCFCFCEAGYRLDEDRGCTSEESVVDLGRUPFYRPLPHIAVLDELPLRFPDDYGAEE 496
QY      401 -----CGD---TNECSI-----NNGG----- 415
Db      497 EAAAEALRGEHTLTKFVCLDSHFGHDCSLTCDGCRNGGTCFPGQDGDCCPEGWTGICN 556
QY      416 QVC-VNTVG---SYEQCHPGYKJHMKKDCVEVKGL--LPTSVSPVSLR---CGKSG 465
Db      557 ETCPPTTFKNCSSPCTCQ-----NGCTCDPLVILGACRCPPGVS---GAHCEDSCP 606
QY      466 GGGCCLRCHSGIH::SSDVT::RTSVTFKLNKGC-SLKNALFPEGURPALPEKHSSYK 524
Db      607 YGKHKRKKCHCA-----NRGRCHRYGACLCDPGLY----- 638
QY      525 ESFRYVNLTCSSGKQVAGAPGRPSTPKEMFIVFELETNQKEVTASCOLSIVKRTKR 584
Db      639 ---RPHLACPPWAGFG----- 660
QY      585 LRKAIRTLKAVHREGFHLQLSOMNLDVAKKPPRTSERQAESCGVCGQHAENQC-VSCRA 643
Db      661 -----EQSH-----TRSCNPKDGSCKKAGFCQSERCAECES 692
QY      644 GTYYGARERCIL-----CPKSTF-CNREGQVTC--- 671
Db      693 GFGPGCRHRCCTCFQVACDPVSGECRTQCQPPGYGEGDGGCECPVGTGTVNCGSGSCV 752
QY      672 EPCPR-----PNSGALKJTPKAWN::XSECGLCQPGEYSADGAPOLCALG-TFOPE 722
Db      753 APCHRVYTGELCFPPGKTG-----EDGADCPBGRKSG::GCQEI::CPACEHGASCP 802
QY      723 AGTSFCPCGGGLATKHQCATSF---QDCETRVQCSPPGHFYNTTHRCI----- 768
Db      803 TGTCLCLP--GFVGSRCQPTCSAGYGTGCTQIRACANDGHCDPTTGRCSAPGWTG::SC 860
QY      769 --ACPVTGYQPE-----FGKNNCVSCPG 789
Db      861 QRACDSGHWGPDCLHPCNCSAGHKRCDVSG 891

RESULT 3
A541C5
fibrillin-2 precursor - human
C:Species: Homo sapiens (man)
C:Date: 05-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 02-Aug-2002
C:Accession: A541C5; S17063; S311C1
R:Zhang, H.; Apfelroth, S.D.; Hu, W.; Davis, E.C.; Sanguinetti, G.; Bonadio, J.; Necham
C:Cell Bio. 124, 955-963, 1994
A:Title: Structure and expression of fibrillin-2, a novel microfibrillar component pre
A:Reference number: A541C5; MUID:94165150; PMID:9:201C5

```

QY	45	DVDECAAGGLDDCHADALCQNTPTSYSKCKXPYQY--EGHQCEDIDECQNLNGCGVHDC	102
Db	1199	DINECSLS-DNLCRNGCVNMIPTYQSCNPGYQATPDRGCTDDEC-MIMNGGDCQQ	1256
QY	103	LNI'PGNYRCTCPGFMLAGHGNCLDVECLENN-----GGQHTCVNVMGSEYECCKEG	157
Db	1257	TNSEGYECSSEGYALPMDGRSCADIDEC-ENNPOTCDGG--OQTNI'PGEYRCLCYDG	1312
QY	158	FFLSDNQHTCIHSEBGLSVNKHGCSHICK---EAPRGSVACGCPGFELAKQORCC	213
Db	1313	FMASMNKTCIDVNECDLN-----SNICMGECENTGSGFICHCQGYVAKGTTGC	1364
QY	214	--ILTNHNGGCG--QHSCDDTADGPSCSCHPYQMTDGRSCLEREDTVLEVTESNTTS	269
Db	1365	TDVDECEIGHNCDMHASCLNIPGSPKCSREGH--IGNIKICIDDEC-----SNGT-	1415
QY	270	VWZGKVRKRRLLMETCAVNNGGCDRTCKDTGTGHCSCPVGFTJLDGKTCKXDIDECQT	329
Db	1416	-----HQCSIN-----AQVNTPGSYRCACSEGT--GDGFTCSDVDECAE	1454
QY	330	RNGGCDH-PCNIVNGSPDCGCKGFKLLTBKSCQDVDECSLBTCT-DHSCINHPGTFFAC	387
Db	1455	NINLCNGQCLNPGVARGCEMGFTPASDSRSCQDIDECFQNICVSGTCNLPGMFHC	1514
QY	388	ACNRGYTL-YGFTHCGDTNCS--IN--NGGCOQCVNTVSGYECQCHPYKLNWKKDC	442
Db	1515	ICDDGYELDRTGNCCTDIDECADPINCWG-----LCVNTGPRVECNPPDFQNLNPTGVGC	1570
QY	443	VEYK-GLLPTSVPGRVLHCKSGGGGCLFRCHSGIHLSSDVTITRTSVTFKLNEGKCS	501
Db	1571	VDRNVCNLYLKFGPR-----GDGS-LSCNTEI---GVGVSRSCC-----CS	1608
QY	502	LKNAELPPEGLRALPEKHSVKEFVYNLTCSSGKQVPCAPGRSTPKEMFITVEFEL	561
Db	1609	LGAWGNPCETCPV-----NSTEYTYL-----CPGEGEGRNP---IT-ILED	1649
QY	562	ETNQKEVTASCD-LSCIVKBTETKRLKAIITLKVAVHREQPHQLSGMNLDVAKKPRTS	620

Qy	45	DVDECAAGLDCEHADALCQNTPTSYKCSKPKVQG--EGRCQEDIDECGMELNGGCVHDC	102
Db	1193	DIINECSLS-DNLCRNGKCNVMIGTYOCSNPGVQATPDQGGCTDIDEC-NMNGGCDTQC	1250
Qy	103	LNT-PGNVRCCTCFQGFMLADHGHNKCLDWDECLENN-----GGCQHTCVVWVGSEYCCCKEG	157
Db	1251	TNSEGSYECSCEGYALMPGRSCADIDEC-ENNPIDCDGG---QCTNIPGEYRCLCYDG	1306
Qy	158	FFLSQDQHTCIHRSEGLSCMNKHGCSHICK-----EAPRGSVACGRPFPELAKNQRDC	213
Db	1307	FMASMDXKTCIDVNECDLN-----PNICYGEGENTKGSFICHCQGYSVKKGTTGC	1358
Qy	214	--ILTCNHNGGC--QHSDDDTADGPSCCHPYKMHETDGRSCLERD'TVLEVTESN'TTS	269
Db	1359	TV'DDECEIGHNCCMHASCLNVPGSFKCSCEGWV--GNGIKCIDJDEC-----ANGT-14C9	
Qy	270	VVDGDKVRKRLIYETCAVNNGGCDRTCKDTSTGVVHCSCFVGTTLQDCKTKCIDDECQT	323
Db	1410	-----HCCS-N-----AQCVNTPGSYRCAGSEGT--GGGFTCSVDVECAE	1448



```

QY      330  RGGCGDH-FCKNIVGSDGCKGKFKLLTDBKSCQDVDECSLRTCDH-SCINHPGTFAC 387
DB      1449  NTNLCENGQC-LNPGAYRCECEMGFTPASDSRSCQDIDECSPQNICVFGTCNN-LPGMFHC 1508
QY      388  ACNRGVTLL-YGTHCGDTNECS--IN--NGGQQQCVVNTVGSYEQCHPGYKLLHNNKQOC 442
DB      1509  ICDDGYELDRTOGNTDIDECADPINCNG--LCVNTPRYECNCPFPCLNATGVC 1564
QY      443  VEVK-GLLPSTVSPVSLHCGKSGGDCGFCRHSGLHSSDVTTIRSVTFKLNEKCS 501
DB      1565  VZNRVNCYLKFGPR-----GDGS-LSCNT---EAGVGVSRSOC-----CS 1662
QY      502  LKNAELFP-EGRLPALPEKHSSVKESPRYNLTCSGKQVPGAGRPSTPKEMFITVEFE 560
DB      1603  LKAGNPNCTCPPNVSHTYTL-----CPGGEGRPNP-----ITILE 1642
QY      561  LETNOKEVTASCD-LSCIVKRTKRLKRAIRTLKRAVHREQPHQLSGMNLVAKKPPRT 619
DB      1643  DIDECEQLPGLCGGNCI----- 1660
QY      620  SERQBSQVGQGH--AE-----NQCVS-----CRAGTYDGAARECIIICP----- 658
DB      1661  NTFGSFQCECPQGYLSEBTRICEIDIECFAPGVGPGTCVNTLGNVTCICPPEYMOVN 1720
QY      659  -----NG-TFONE-----EGQMTCEPCPRPKSG- 681
DB      1721  GGHNCMDMRKSCFYRSYNGTTCEBELPNVYKRMCCCTYNVKAGNKCPCPCPTGTADF 1780
QY      682  -----ALKTPAMNSEC----- 694
DB      1781  KTCIGNIPGFTDIHTGRAVDICEKEIPGICANGVCINQIGSPCECPTGFSYNDLLV 1840
QY      695  -----GGLCO-----PEYSADGAPQCALG-TFQPEAG---RTSCPP-- 730
DB      1841  CEDIDECNSGNDLCORNADCNPSGYRCE-----CAAGFKLSPNAGCVDRECLEIP 1893
QY      731  --CGGGLATKHGATSFQ-DCETRVCQSPGFYNTTHRCIR--CPVGTQYQPEGIONCV 785
DB      1894  NVCSHGLCVDLQG--SYQICLNNGFKASQDQYCMQDVDECEHPCANGCTCKNTVGSYNCL 1951
QY      786  SPCG-----NTTTFDGSNTI--TQCKNRRCGGELGDF-----TGYTESPNYFGNVPAN 832
DB      1952  CYPGPELTHNNDCLDIDECSSFFGVCRNGRCFNEIGSFKLCNNEGVELTDPGKNCIDTN 2011
QY      833  TEC 835
DB      2012  EC 2013

RESULT 5
A:55567
C:Species: Bos primigenius taurus (cattle)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Aug-2002
C:Accession: A55567
R:Filistra, D.J.; Li, L.; Potter, K.A.; Womack, J.; Byers, P.H.
Genomics 23, 480-485, 1994
A:Title: Sequence of the coding region of the bovine fibrillin cDNA and localization to
A:Reference number: A55567; MUID:95137597; PMID:7835900
A:Accession: A55567
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2871 <TIL>
A:Cross-references: GB:L28748; NID:9508427; PIDN:AAA74122.1; PID:9508428
A:Superfamily: fibrillin 1; EGF homology
F:1201-1236/Domain: EGF homology <EGF>

Query Match      12.3%; Score 691; DB 2; Length 2871;
Best Local Similarity 22.7%; Pred. No. 7,6e-34;
Matches 271; Conservative 107; Mismatches 375; Indels 440; Gaps 49;

QY      45  PVDECAQGLDDCHADALCONTPTSYKCSCKPGYQGEGRCCEDIDECNELNGGCVHDCLN 104

```

```

DB      1363  DLEBSGTHMSCADCKNTMGSYAC-CKEGYTGGTCTDCLDECSNMLNLCNGGQCLN 1422
QY      105  IPGNVACTCFQGFMLAHGHNCLDVEEC-----LENNG 137
DB      1423  AFGYRCECDMEFVPSADGACEDIDECSLPNICVFGTCHNLPGLPFCEBISGVELDSSG 1482
QY      138  G-----CQHTCUNWGSYECCKEGFFLSNQHJTCIH-----RS 171
DB      1483  GNCIDWECDDPTTCISGNCVNTPSYTCDCPPDFELNPTRVGCVDTSSGNYLDIRPRG 1542
QY      172  BEG-LSCMN-----KDHGCSHI----- 187
DB      1543  DNGDTACSNE-LGVYSKASCSCSLKAKTPELCPPVNTSEYKILCPGGEGFRNPITV 1602
QY      188  -----CKEAPR-----GSVACEAPGFELAKNORDC--LTCNH-----GNNG 223
DB      1603  LLEDIDECQELPGLCGGKGCINTFGSFQCRPTGYVLENTDTRVCDVNECETPGICGEG- 1661
QY      224  CQHSDDTADGPECSCHPOYKMHITGERSCLEREDTVLEVTESNTTSVAVDGGKRVKRLLM 283
DB      1662  ---TCVNTVGNVTCICPPDYKQVNGNNDNRSLCYRNYADNQTCDGELLFMYTKKM 1718
QY      284  ETCVAVNG-GCDRTCK-----DTSTGV----- 304
DB      1719  CCSYNIGRAMNKPCEQCPISTDEPATLGSQRPGFVIDIVTG-LPVDICEKEIFGVE 1778
QY      305  -----HCSGVGFTLQDKTKCDIDEC----- 328
DB      1779  NGVCINMVGSRCEPVGFFYNDKLLVCEIDECQNGPVCORNAECINTAGSYRCDCKPG 1838
QY      329  ---TRNGGCDH-----PCXNI-VGSFGCKGKFKLLTDBKSCQDVDECSLD 371
DB      1839  YAFSTGQCNDRNECQELFNISSHGCCIDTVGSFVLCHTGFKTNADOTMCLDINECEPD 1898
QY      372  RTCDHSCINHPGTFAACNRGYTYLGFTHCGTNECSINNGG-CQQ-CVNTVGSYEQC 429
DB      1899  ACNGTGRNTIGSFNCRNNGHFIILSHNNDICIDVDECATGNLNCNGQICINTVGSFQCC 1958
QY      430  HGYKLLHNNKCKVEVKGLJ---PTSVSPRVS-HCKKSGGDCGF-PCHSGLHLSSDVTI 487
DB      1959  NEGYEVAPGRTCVQINECLLPKCAPOGT---CNLDGSYRCI--CPFGVSLQDKCE- 2012
QY      488  R-SVTEKLE-----GKCSLKNAE-----JFPEGLR-PALPEKHSSVKESPRVYN 531
DB      2013  -----DIDECVEBPEICALGTC--NTEGSKCLCPGFSLSSTGRCDLRMSYCYAK 2064
QY      532  LTCSSKQVPGAPRPSTPKEMFITVEPELTNQKEVTASCDLS-----CIVKATEKR 584
DB      2065  F-----EGGKCSPK-----SRNHSKQECCLXGEGMGDPCELCPTED- 2103
QY      585  LAKAIRTLKRAVHRECFH-LQL-----SGXKLDVAKKPERTSERQAES----- 626
DB      2104  -----PEAPRQICPYSGIIVGDDSDAVDMDECKEPDVCCKHGGC:WTDSSY 2149
QY      627  ---CGVGQGHAEQCV---SCRAGTYDGAARECICPNNGTFSQNEEGMTCPCPRPQNS 680
DB      2150  RCSCPFGYILQNECVTDECSVGN-----PCGKGTCKNVIGGFEC-TCEEGEP 2198
QY      681  GALKTPENANMSECG---LCQPGEYSDGAPAPQCALGTGTFQPEAGRISC---PFCGGG 734
DB      2199  GPWMTCE--DINECAONPLLCAPRCVNTYGSYECK-CPAGYVYLR-DRNMCKDEDECEG 2254
QY      735  LAKHOGATSFQDCETRVQ-----CSPGFYNTTHRCICRCPVGTYQPEFGKNQCVSPG 789
DB      2255  ---KHDCAEKQNECKULICTYLICIGPGYQRPDPDEGCV-----DENECTQAPG 2300
QY      790  NITTFDGSNTITQCKNRRCGGELGDFT-----GYIESPNYFGNYPANTECWTINPFPK 844
DB      2301  -----ICENGRCLNTRSYTCECNDGTFASN----- 2327
QY      845  RRLI:VPEIFILPDEDCG-----YLVNRKTSSSNSVTTTETCTQTYER----- 888

```

Db 2328 -----QDECLNREGYCFTEVLQMCQIGSSNRPVTKSECCCGRGWGP 2373

Qy 889 -----PIAFTSRSKLWQFKSNEGSARGFQVPVTVYDEYQELIEDIVRDG 936

Db 2374 HCEICPFQQTAFKLC-----PHGRGFMTNGADICE--CKVIHVDVCRNG 2416

RESULT 6

A47221

fibrillin 1 precursor - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 02-Jun-1995 #sequence revision 25-Apr-1997 #text change 02-Aug-2002

C:Accession: A47221; 154355; S17064; 159574; S17062; S62111; A34198

R:Corson, G.M.; Chalberg, S.C.; Dietz, H.C.; Charbonneau, N.L.; Sakai, L.Y.

Genomics 17, 476-484, 1993

A:Title: Fibrillin binds calcium and is coded by cDNAs that reveal a multidomain structure

A:Reference number: A47221; MUID:94010947; PMID:7691719

A:Accession: A47221

A:Molecule type: mRNA

A:Residues: 1-337, 'T', 339-1029 <COR>

A:CROSS-references: GB:X63556

R:Pereira, L.V.; D'Alessio, M.; Ramirez, F.; Lynch, J.; Sykes, B.; Pangilinan, T.; Bonad

Hum. Mol. Genet. 2, 961-968, 1993

A:Title: Genomic organization of the sequence coding for fibrillin, the defective gene p

A:Reference number: 154355; MUID:93372860; PMID:8364578

A:Accession: 154355

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 132-3202 <PER>

A:CROSS-references: GB:LI13923; NID:9306745; PIDN:AAB2036.1; PID:9306746

R:Maslen, C.L.; Corson, G.M.; Maddox, B.K.; Glanville, R.W.; Sakai, L.Y.

Nature 352, 334-337, 1991

A:Title: Partial sequence of a candidate gene for the Marfan syndrome.

A:Reference number: S17064; MUID:91304568; PMID:1852207

A:Accession: S17064

A:Molecule type: mRNA

A:Residues: 1030-3002 <MAS>

A:CROSS-references: EMBL:X63556

R:Dietz, H.C.; Valle, D.; Francomano, C.A.; Kendzior, R.J.

Science 259, 680-683, 1993

A:Title: The skipping of constitutive exons in vivo induced by nonsense mutations.

A:Reference number: 159574; MUID:93157831; PMID:8430317

A:Accession: 159574

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 2217-2288, 'I', 2290-2325 <RES>

A:CROSS-references: GB:S54426; NID:9264860; PIDN:AAB25244.1; PID:9264861

R:Lee, B.; Godfrey, M.; Vitale, E.; Horii, H.; Mattei, M.G.; Safarazi, M.; Tsipouras, P.

Nature 352, 330-334, 1991

A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two differ

A:Reference number: S17062; MUID:91304567; PMID:1852206

A:Accession: S17062

A:Molecule type: mRNA

A:Residues: VLVTVVFIFSYNKL', 944-1444 <LEE1>

A:CROSS-references: EMBL:X62008; NID:931398; PIDN:CAB56534.1; PID:95924015

A:Accession: S62111

A:Molecule type: protein

A:Residues: 1166-1176, 'X', 1178-1180, 'D', 1182-1185 <LEE2>

R:Maddox, B.K.; Sakai, L.Y.; Keene, D.R.; Glanville, R.W.

J. Biol. Chem. 264, 21381-21385, 1989

A:Title: Connective tissue microfibrils. Isolation and characterization of three large f

A:Reference number: A34198; MUID:90078246; PMID:2512293

A:Accession: A34198

A:Molecule type: protein

A:Residues: 565-575, 1890-1892, 'I', 1894-1900 <MAD>

C:Comment: Fibrillin is a major component of elastin-associated microfibrils.

C:Genetics:

A:Gene: GDB:FBN1

A:CROSS-references: GDB:127115; OMIM:134797; OMIM:154700

A:Map position: 15q21.1-15q21.1

A:Introns: 2236/1; 2258/1; 2297/1

C:Superfamily: fibrillin 1; EGF homology

C:Keywords: alternative splicing; calcium binding; extracellular matrix; glycoprotein; M

F:13062/Product: fibrillin (5'-region exon A splice form) (fragment) #status predicted  
 F:132-3002/Product: fibrillin (5'-region exon C splice form) #status predicted <MAD>  
 F:1332-1367/Domain: EGF homology <EGF>  
 F:1457-1492/Domain: EGF homology <EGF2>  
 F:2262-2295/Domain: EGF homology <EGF1>

Query Match 12.3% Score 691; DB 2; Length 3002;

Best Local Similarity 22.5% Pred. No. 8e-34;  
 Matches 269; Conservative 105; Mismatches 378; Indels 442; Gaps 45;

Qy 45 DVDECAOGLDCHADAJCCQNTPTSYKSCCKPGYQGEGRQCEDIDEGCNELNGGCVHDCLN 164

Db 1494 DLDECSNGTMCQSHADCKVTMSYRCLCKEYTGDTCTDDESENKLCGNGOCLN 1553

Qy 105 IPGNRYCTCFDGYLAHDHNCLEVDECSLNNGGCQHTCVNMGSYECCKEYFFLSDNQ 164

Db 1554 APGYRCECMGFSPADGKACEDIDECSLPNICVFCTCHNLPLGRCECEIGVELDRSG 1613

Qy 165 HTCHRSSEGLSCNKHGCSHICKAPROSVACEGPFELAKNQRDCILT----- 216

Db 1614 GNCTDYNE-----CLDPTTCISGNCVATP-GSY-CDCPPDFELNPTRVGCVDTREGCYLD 1668

Qy 217 ----- 216

Db 1669 IRPRGNGDFACNEIOGVSKASCCSLKAKWTPCEMCAVNTSEYKLCRGGEGFRP 1728

Qy 217 -----CNHSGNGGCG-----HSCDDCTAD----- 233

Db 1729 NPITVILEDDICQELSLGCGGGKICNTFGSFQCRCPYGVYLNEDTRVCDVNECEPGL 1788

Qy 234 -GP-----ECSCHPOYKHTOGRSCLEREDTVLEVTESVTVSGDKAVKRLJ 282

Db 1789 CGPOTCVTYGNTYCTICPPDYMQVNGGNCMDRRSLCYRYADNCTCGELLFNMTXK 1848

Qy 263 YETCAVANG-GC9RTCK-----DTSTGV----- 304

Db 1849 MCCSYNIGRAMNKPCECPIPTDEPATLCGSGRPGFVIDYVGLPVIDEGREIPGVC 1908

Qy 305 -----HGCSPVGFLLQDGRKTKDIDECQ----- 328

Db 1909 ENGVCINWVSFRCECPGVFFYNDKLLVCEIDECQNGPVCQNAECINTAGSYRCCKP 1968

Qy 329 -----TRNGGCDH-----PCNIVSGFDGCKKGFKLLTDEKSCQDVDECSL 370

Db 1369 GYRFTSTGCKDRNECOEINICSHGQCIDTVGSFVCLHTGKNTDCTMCLDINCEER 2028

Qy 371 DETCDHSCINHPGTFACACVNGVTLVGFTHCGTNECSINNGG-CQQ-VCVNTVGSYECC 428

Db 2029 DACNGTCRNTIGSFNCRCHGFFLSHNNDCLDVDECAISGNLGRNGQCINTVGSFQCC 2088

Qy 429 CHPGYKLFHWKDKCVYKGLL--PTSVSRVSLHCGKSGGGGCGFLRCHSGIHLSDDVT 486

Db 2089 CNEGYEAPDORTCVINECLLEPRKCAPGT---CONLDGSYRCI---CPGYSLSQNEKE 2143

Qy 487 IRTSVTKLANE-----GKCSLQAE-----LPPEGLR-PALPEKISSVKEFRIV 530

Db 2144 -----DIDECVEEPICALGTCS--NTSGSFKCLCPGEGFSLSSSGRRQCDLRMSCYA 2194

Qy 531 NLTCSGKQVPGAPRESTPKEMFITVEFELETNKEVTASCDLS-----GIVKRETEK 583

Db 2195 KF-----EGKCSSPK-----SRNHSKQECCLAKGSGWSDPFCBLCTEP 2234

Qy 584 RLRKAIRTLKAVHREQFHQL-----SQNLDVAKKPPRTSERQAES----- 626

Db 2235 -----DEAFQICPYGSGIIVGPDSDAVDVECKEPVCKHGQCCINTDGS 2279

Qy 627 ----CGVQGHAEQCV---SCRAGTYDGAERECILCPNGTTPONEGQVTCPCPRPGN 679

Db 2280 YRCECPFGYTLAGNECVDTDECSVGN-----PCGNGTCKNV-GGFEC-TCEEGFE 2328

Qy 680 SGALKTFEAMNMSECG-----JCQGEYSADGAPCQICAJCTFQEGAGRTSC---FFCGG 733

Db 2329 PGPMWTCE--CINECAQNPLCAFRVNTYGSYECK-CPVGIVLFE--DREMCKDEJCEE 2394

```

QY 734 GLATKHQGNATSFQDCETRVQ-----CSPGHFVNTTTHRCIRCFVGTQYQPERGKNNVSCVP 788
DB 2385 G---KHDCTEKQMECKNLTGYTMCICGFGYQRRPQGEQV-----DENECQTKP 2430
QY 789 GNTTTFDQSTNITCKNNRRCGELGDF-----CYIESPNYPGNYPANTBCTWTIRPPP 843
DB 2431 G-----ICENGRCLNTGSYTCRCNDGFTASPN----- 2456
QY 844 KRRILVWPEIFLP:EDDCGD-----YLWKRKTSSSNVTTTTCQTYER--- 888
DB 2459 -----QDECLDNREGYCFTEVLQNMCIQSSNRNPNVTXSECCCGGRGWG 2503
QY 889 -----PIAFTSRKSLWTFKSNRGNARSARGQVPPVYDYDEYQBELIEDIVRDG 936
DB 2504 PHCEICPFGTVAFKKLC-----PHGRGFMINGADIDE--CKVIHDVCRNG 2547

RESULT 7
A55624
Fibrillin-1 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 02-Aug-2002
C:Accession: A55624
R:Vin, W.; Smiley, E.; Germlinger, J.; Sanguinetti, C.; Lawton, T.; Pereira, L.; Ramirez,
J. Biol. Chem. 270, 1798-1806, 1995
A:Title: Primary structure and developmental expression of Fbn-1, the mouse fibrillin gene
A:Reference number: A55624; MUID:95130561; PMID:7829516
A:Accession: A55624
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2871 <YIN>
A:Cross-references: GB:L29454; NID:9575509; PIDN:AAA56840.1; P:ID:9575510
C:Genetics:
A:Gene: Fbn-1
A:Superfamily: fibrillin 1; EGF homology
F:1201-1236/Domain: EGF homology <EGF>

Query Match 12.3%; Score 690; DB 2; Length 2871;
Best Local Similarity 23.4%; Pred. No. 8.8e-34;
Matches 276; Conservative 104; Mismatches 384; Indels 418; Gaps 51;

QY 45 DVDECAQGLDDCHADALCONTPTSFKCSCKPGYQGEGRQCEIDECGNELMNGCVHDCLN 104
DB 1363 DIDECSNGTHMCSQADCKNTWGSYRCLCKDGTGGFTCTDLDCSENLKLCGNQCCLN 1422
QY 1205 IFGNVRCCTCFDGFMLAHQHNCLDVDEC-----JENNG 137
DB 1423 APGGYRCEDMGFVPSADGKACEDIDECSLPNICVFGTCHNLPLFRCECEBIVGELDRSG 1482
QY 138 G-----CQHTCVNMVGSYCCCKEGFFLSDNQHCTCIH-----RS 171
DB 1483 GNCVTQNECLDPTTCISGNCVNTPGSYTCDSPDFELAPTRVGCVDTRSGNLYNIRPG 1542
QY 172 BEGLSOMKDHG-----CSHI----- 187
DB 1543 DNGDTACSNIEIGVGSKASCCSLGKAWGTPCELCSVNTSEYKILCPGGEGRFNPIIV 1602
QY 188 -----CKEAPR-----GSVACECPRGFEJAKNQDCC--ILTCNH-----GNNG 223
DB 1603 ILIEDICQELPGLCGGKCIINTFGSFQCRQPTGYVLNEDTRVCDVNECECTPGICGPG- 1661
QY 224 CQHSCTDADGPECSCHPOYKMTDGRSCLEREDTVLEVTESNTTSSVVDGDKVRKRLJM 283
DB 1662 ----TCVNTGVNTYICFPDYMVGNGNCKMDRRSICRYNYADNCTCGELLFNMTKKN 1718
QY 284 ETCAVN-----NGCORTCKDTST----- 302
DB 1719 CCSSYNIRRAWNKPCEQ-CPIPTDDEFATLGSQRPGFVIDIYTLGLPVDIDECREIPGVC 1777
QY 303 --GV-----HCSVPVGTQLDQKTKCDIDECQ----- 328
DB 1778 ENGVCINMVGSRCECPVGFVFNKDLLVGEDIDECQNGPVCLRNAECINTAGSYRCDCCKP 1837

```

```

QY 329 ----TRNGCCH-----FCNIVGSEFDCCGCKHGFKLLTDEKSCQCVDECSL 370
DB 1836 GYRLTSTGQCNRNECQEIPINICSHGQCIDTVGSFVCLHTGFKTNEQTVCLDNESE 1897
QY 371 DRFCDSHSCINHPGTPACACNRGYTLVGFHCQGTNECSINNG-CQQ-VCVNTVGSYEQ 428
DB 1998 DACNGTCRTNTIGSFYRCRCNHGFIILSHNNDCIDVDCATGNGNLCRNGOCVNTVGSFQR 1957
QY 429 CHPGYKLRNKKDCVEVKG--LTPSVSPRSLHCGKSGGDCGFLRHSGLHLSSDYTT 486
DB 1958 CNEGYEVAEDGRTCYDINECVLDPGKCAPGT---CQNLDSYRC--CPPGYSLQNDKCE 2012
QY 487 IRTSVTFKNE-----OKCSKNAE-----LFPEGLR-PALPEKHSVSKSPRYV 530
DB 2013 -----DIDECVEEPEICAJGTCS--NTEGSFKCLCEGFSWSSSGRRCCQDRMSYCYA 2063
QY 531 NLTCSGKQVPGAPGRPSTPKEMFITVFELETNKEVITASCDLS-----CIVKRTEK 583
DB 2064 KP-----EGKCSSPK-----SRNHSKQECCLALKEGWDGDPCELCPTEP 2103
QY 584 RLKKAIRTLRKAVHREQFHLQ-----SGMNLQVAKKPP-----RTSER 622
DB 2104 -----DEAPRQICPFGSGIIVGPDSDAVDMDECKEPDVCRHGQCINTQGS 2148
QY 623 CAESCGVGQGHAEQCV---SCRAGTYDGAERCLCPNGTFFQNEEGOMTCPCPRPN 679
DB 2149 YRCECPFGVILEGNECVDTDECSVGN-----PCGNGTCKXVIGGFEC--DEGEFE 2197
QY 680 SGALKTPPEANMSECGG---LQCPGEYSADGFAPQCALGTQFQPEAGRTSC---FPCGG 733
DB 2198 PGPMTCE--DINECAQNPJLCAFRCVNTYGSYECK-CPGVYVJRE-DARMCKDEDECAE 2253
QY 734 GLATKHQGNATSFQDCETRVQ-----CSPGHFVNTTTHRCIRCFVGTQYQPERGKNNVSCVP 788
DB 2254 G---KHDCTEKQMECKNLTGYTMCICGFGYQRRPQGEQI-----DENECQTKP 2299
QY 789 GNTTTFDQSTNITCKNNRRCGELGDF-----CYIESPNYPGNYPANTBCTWTINIRPP 843
DB 2300 G-----ICENGRCLNTGSYTCRCNDGFTAS-----PTODECL----- 2332
QY 844 KRRILVWPEIFLP:EDDCGDYLWKRKTSSSNVTTTTCQTYER-----PIAFTS 894
DB 2333 DNEGYCFSEVP--ENMC-----QGSNRPVNTXSECCCGGRGNGLHCEICPEGTV 2384
QY 895 RSXKLWIOQKNEGNSARGQVPPVYDYDEYQBELIEDIVRDG 936
DB 2385 AYKKLC-----PHGRGFMINGADIDE--CKVIHDVCRNG 2416

RESULT 8
T46488
Hypothetical protein DKFZp434J065.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C:Accession: T46488
R:Dueserthoeft, A.; Lauber, J.; Yewes, H.W.; Gassenhuber, J.; Wienann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23035
A:Accession: T46488
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-741 <AAA>
A:Cross-references: EMBL:AJ237638
A:Experimental source: adult testis; clone DKFZp434J065
C:Genetics:
A:Note: DKFZp434J065.1

Query Match 12.3%; Score 687; DB 2; Length 741;
Best Local Similarity 31.7%; Pred. No. 3.3e-34;
Matches 155; Conservative 61; Mismatches 196; Indels 70; Gaps 15;

QY 37 GRAAGPQEDVE--CAQGLDCHADALCONTPTSFKCSCKPGY--QGEGRQCEIDECGN 92

```



QY 45 DVDECAQGLDDCHADALCONTPTSYKCSCKPGY--QGEGRCEIDIDE-----GNEL 94  
Db 907 DIDECAQAQHLCL-SQRCENTEGSLCLCPAGFIASEEGSNCLDVCLRLRDEVCGRICI 965  
QY 95 N-----GGC-----VHDCNLIPGNRYCT-CFDFGFMIAH :21  
Db 966 NTAGAFRCBYCDGYRMSRRGHCEIDIELTPSTCPESQCVNSPGSYQCVPTCEGR--- 1022  
QY 122 DGHN--CLDVDECLNNGCOHTCVNVGSGYECCKKEGFFLSLDRKHTC--IHRSEGLSC 177  
Db 1023 -GWNQGLDVDECLQVKVCTNSGCTNLGSGYMCCHKGYSTPPDRHRHQDIDECQCNLC 1081  
QY 178 MNKDHGCHSHCKEAPRGSVACBRCRPFEL--AKNQDCLILCNHGNGCOH--SCDITADG 234  
Db 1082 MNGQ-----CKNTD-GSFRCTGGGYQLSAKQDCEIDICEHRH-LCSHGQCRNTEGS 1133  
QY 235 PGSCCHPQYKMTDGRSLEREDTIVLESTESNTTSVVDGDKRVKRLMETCAVNNGCCD 294  
Db 1134 FCLCNQGYRASVLGDHC---ED-----INE-----CLEDSVYCCQG-- 1167  
QY 295 RTCKDTSTGWHSCPCVPGFTLQDDGKTCKDIDECOTRNGGCDHF--CKNIVGSFDCGCKKG 352  
Db 1168 -DCINTAGSYDCTCDGLQLN-DNKGQDINEC-AQPLGLAPHGEC-LNTQGSFVCEQG 1224  
QY 353 FKLLTDEKSCQDVDECSLDRTCD-HS-CINHPGTTFACACNRGYTL-YGFTCHGDTNECSI 409  
Db 1225 FSIADGRTCEIDIECVNNTVCDSHGFCNDTAGSFRLCYQGFQAPQGGQGVNNECEL 1284  
QY 410 NNGGC-QQCVNVTGSEYEQCHPGYKLR-----WNKXDC--V 443  
Db 1285 LSGVCGEAFCEVNEGSFLCVCADENOESYPMTGQCRSATEDSGVDROPKEKCEYNL 1344  
QY 444 EVKGLLPTSVSRVSLH--CGKSGGGDGCFLACHSGIHLSSDVT\*IRTSVTFKLENGKCS 501  
Db 1345 NDASLCDNVLANVYKQECCTSGAGW-----1372  
QY 502 LKNAELFP-----EGLRALPEKHSSYKESFR-----VNLTCSSGK 538  
Db 1373 -DNCEIFPCPVQGTAFBSEMCPRGKGFVPAGESSEYETGGENYKDADECLFGEIBCKNGY 1431  
QY 539 QVPGAPGRSTPKEMFITEFELETNQEVATASCDLSCLVKEKTEKLRKA\*RTLKRAVHR 598  
Db 1432 CLNTQPGVECYCKE-----GTVPYDKLQC-----1456  
QY 599 EGFHLQSLGNMLDVAKKPRTSERQAESCGVQGHAEHQCVS-----CRAGTYVDEGA 650  
Db 1457 -----FDVDECDP-----NSC-----IDGQCVNTEGYNCFCTHPWLDAS 1483  
QY 651 RERCILCPNGTFCNBEQWTCBPC-----PRGNSGALKT-----PEANNVSE 693  
Db 1494 EXRCVQPTSSNEQIBETDVYQDLCEHLSSEYVCSRPLVGRQTTYTECCCLYGEANGW-Q 1552  
QY 694 CGSLQCPGYSADGAPCOLCALGTFOPEAGRTSCF-----PCGGSLATKHQ 740  
Db 1553 C-ALCPMKD--SDDYA--QLONI-----PVGRRPYGRDALVDPSQGPETDPFFQCR 1603  
QY 741 GATFOCCETRVOCSGPHFYNTTTHRCIRCPVG-----TYQEPFGKNKCVSCFGNTT 793  
Db 1604 FJNSFEELQAE-EC--GILNGCNGRCVRVQEGYTCDCFDGYHLDMAKMTVCQV--NECS 1658  
QY 794 DFDGSTNITCKNRRCGGLGPF-----TGYTES--PNYPGNYPANT 833  
Db 1659 ELNNRMSL--CKNAKINTEGYSKVCULPGYVPSDKPNY--CTPLNT :701

## RESULT 11

A49457  
fibulin-2 precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 08-Sep-2002  
C:Accession: A49457; S74095  
R:Pan, T.C.; Sasaki, T.; Zhang, R.Z.; Faessler, R.; Timpl, R.; Chu, M.L.  
J. Cell. Biol. 123, 1269-1277, 1993

## RESULT 12

A55494  
latent transforming growth factor-beta-binding protein - human.  
C:Species: Homo sapiens (man)  
C:Date: 03-Mar-1995 #sequence\_revision 03-Mar-1995 #text\_change 11-Jan-2000  
C:Accession: A55494  
R:Mooren, A.; Olofsson, A.; Stenman, G.; Sahlin, P.; Kanazaki, T.; Claessens-Welsh, L.; et al.  
J. Biol. Chem. 269, 32469-32478, 1994  
A:Title: Identification and characterization of LTBP-2, a novel latent transforming growth factor-beta-binding protein.  
A:Reference number: A55494; MUID:95096101; PMID:7798246

A:Title: Structure and expression of fibulin-2, a novel extracellular matrix protein w  
A:Reference number: A49457; MUID:94064787; PMID:8245130  
A:Accession: A49457  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1221 <PAN>  
A:Cross-references: GB:X75285; NID:9437046; PIDN:CAA53040.1; PID:9437047  
R:Sasaki, T.; Mann, K.; Yurkiv, G.; Chu, M.L.; Timpl, R.  
Eur. J. Biochem. 240, 427-434, 1996  
A:Title: Different susceptibilities of fibulin-1 and fibulin-2 to cleavage by matrix m  
A:Reference number: S74095; MUID:96439073; PMID:8841408  
A:Accession: S74095  
A:Molecule type: protein  
A:Residues: 236-238 'X', 240-247:260-275:336-344, 'L', 346-361:405-426:566-568, 'E', 569-5  
C:Superfamily: fibulin-2; EGF homology; von Willebrand factor type C repeat homology  
C:Keywords: calcium binding; duplication; extracellular matrix; glycoprotein; hcmotr.f.m  
F:942-978/Domain: EGF homology <EGF>

Query Match 10.3%; Score 577.5; DB 2; Length 1221;  
Best Local Similarity 29.0%; Pred. No. 2.4e-27;  
Matches 157; Conservative 66; Mismatches 182; Indels 137; Gaps 29;

QY 44 EDVDECAQGLDDCHADALCONTPTSYK-----SCKPGY-QSGRQCEIDIECGNELNNG 97  
Db 708 EDQDECLMTHPCSKWKQFCWNTLGSFYCVNHTVJCAEGYILNAHRKCVDINECVTDLH-T 766  
QY 98 C-VHDCNLIPGNRYC---TCFDGFMIAHGHKCLDVDECLNNGGQ--HTCVNVMGS 149  
Db 767 C\*RAEHCNTPSPGFCYKALTCPEGYVLT-DG-ETDVEDCVTGTNCOAGSFCONTYKGS 824  
QY 150 YEC-----CCKEYFFLSDNQHCTIHRSE-----EGLSCNKHGCHSHCKEAPRGSV 196  
Db 825 FYQASQRQMDG-FLQDPGNCVDINECTSLLEPCRSQGFSC-NTVG--SYTCORNP---- 877  
QY 197 ACECRPGFLAKNQDC--ILTCNHNNGGQCHS--CDDTADGPECSCHPQYKMTDGRSC 252  
Db 978 -VCGRGYHANSEGECDVDNECETGVHRCGEGQ-CYNLPGSYRCDCCKPGFORDAFGRTC 936  
QY 253 LREEDTVLEVTESNTTSVVGGSKRVKRLMETCAVNNNGG-CDRTCKDTSTGWHSCFPVG 311  
Db 937 :D-----VNECVNTPGRLCQHTCENTPSSYRSCAAG 968  
QY 312 FTLLQDLGKTKIDECQTRNGGCDHFCKNIVGSFDCGCKKGKFLLTDEKSCQDVDECSLD 371  
Db 969 FLAADGKHECDVNECETSR--CSQECANIVGSYCYCRQGYQLAEDGHTCTDIDECAG 1026  
QY 372 R--TCDSHSGINHPGTFACAC-NRGYTL-Y-GFTHCQDTNECSINNGGQ--QVCNNTVGSY 425  
Db 1027 AG-LCTFCRVNVPVGSYQACPEQGYTGMXANGRSCKDLDECALGTNCSAEETCHNIQGSF 1086  
QY 426 EC---QCHPGY---KJHNKXKDCVSEV-----KGLL-----449  
Db 1087 RCLRFDCFPYVRYVRSQTKCERTTCQDI-ECQTSRPARITHYQLNFQTSLLVPAHIFRIGPA 1146  
QY 450 PTVSVPRLVHCKGSGGDCGF-RCHSGIHLSSDVT\*IRTSVTFKLN--EGKCSXKXEL 507  
Db 1147 PAFAGDTISLTITK-GNEEGYP-----VTRFLNAVTVGWSLQRSVL 1186  
QY 508 FP 509  
Db 1187 EP 1186

A:Accession: A55494  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1820 <MOR>  
A:Cross-references: GB:Z17976  
C:Genetics:  
A:Gene: GDB:LTBP2  
A:Cross-references: GDB:568901  
A:Map position: 11pter-11qter  
C:Superfamily: unassigned EGF-related proteins; EGF homology  
F:1222-1257/Domain: EGF homology <EGF>  
F:1525-1565/Domain: EGF homology <EGF1>

Query Match 10.2%; Score 570; DB 2; Length 1820;  
Best Local Similarity 23.2%; Pred. No. 1e-26;  
Matches 208; Conservative 110; Mismatches 339; Indels 240; Gaps 45;

```
QY 46 VDECAQGLDDCHADALCONTPYSKCKPGVQGEGRQ--CEDIDCEGNELNCGVHD-- 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 845 IDCAAGATNVGPGTCVNLPGYRCVCGSPGVQLHPSOAYCTDNEC---LRDPCQKGR 901

QY 102 CLNIPGNRYRCTCFDGFMLAHDG--HNCLDVDECLNNGGCO-HTCVNVMGSECCCKECP 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 902 CLNIRVGSFCFCYPTGTTATSGATQECQDINEC-EQPGVCGGQCTNTEGSHCECDGG 960

QY 159 FL-----SDNQHCTCI-----HRSEGLSCMKND----- 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 961 IMVRKGHCQDINECRHPTGCPDGRCVNSPGSYTCLACEEYRGQSS-SCVDVNECLTGV 1019

QY 182 --HG--CSHICKEAPRGSVACRPGFELAKNORDCILTCNHNCGGQCHSDDTADGPEC- 237
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1020 CAHKGKTNL-----EGSRFCSCEQYEYVTSDEK-----GCQ-DVDECASTRSCP 1062

QY 238 -----SCHPOYKMHDTGRSLREDTVLEVTESNT--SVVDGSKRVKGR-LJM 283
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1063 TGLCLNTEGSPACSCACNGYVNEBGTAC-----ED-----L 1094

QY 284 ETCVANNGGCDRTCKDTSTGVHC-SCPVGFTLQDQKTKDIDECQTRNGGC-DHFCRN- 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1095 DECAFPVCPGVCNTAGSPFCKDGGYRSPLDGSDCEDVDECDPQSSCLGSECNT 1154

QY 342 VSGPFGCKKFKLLTDEKSCQDVDECSLDRTC--DHSCINHPGTFACNNGY--TLVGF 398
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1155 VGSYQCLCPQGFQ--LANGTVCEVNECEMGEBCAPHGECCLASHGSPFCLCAPGFVSARG 1213

QY 399 THCGDTNECSINNGCQCVNTVSGYECQCHPGYKLNHNKDCVEYKGLLPVSFVRS 458
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1214 TSCQDVDECATTPCVGCHVNTGSEFNCLCEFGQPSPESECEVDIDECEDYDGPVCGT 1273

QY 459 LHCXSGGGDGFRLCHGSHLSSDVTITRTSVTFKLNKGCCLKNAELFPEGLRPAALPE 518
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1274 WRCENSPGSYRCVLGCPQGFHMAPNGDCI--DIDECANDTMCG----- 1314

QY 519 KH---SSVKESEFYNLTSSCKQVPGAPGRSTPKEMFIVPEFL-----ETNOKE 567
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1315 SHGFQDNTGSRFL---CDQG-----FEISPSGMDVCDVNECE 1350

QY 568 VTASCDLSICVIRKTKRLKARTLRKAVHRQFHLQLSGMNLVAKKPPR--TSEROAE 625
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1351 LMLAVCGAALCENVEGSLCLCASOLEBYDAGEGTAHAGLEVRLVCLRPQDRHAPATRM 1410

QY 626 SCGVGO-GHA-----ENQVCSCAGTYPDGARERCILCPNGTFQNEBGMQTECP 673
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1411 DCYSGKGHAPCSSVLGRNTTQAECCCTQGTATGDA-----CDLCP-----SEDSAEFSEI 1461

QY 674 CRPGNSGALKTPEAW-----NMSEC-----GGLCQPGE--YSADGFAPCOLCALGTF 719
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1462 C--PSKGYPVEGAWTFQGTWYTDACVIFGPGLCPNRCNLNTPGVV-C-LCNPQ-F 1516

QY 720 QPEAGRTSCPPGGGLATKHQGATSPQDCETRVQCSPGHFYNT--THRCIRCPVGTIYQPE 778
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1517 HYDASHKKC-----EHDQEQ--JACENGEVNTGSEFHCFCSPFLTJ--D 1559
```

QY 779 FGKNCVSCPGNTTDFDGSNTITQCKNRRCGGELGDFDTGYIESPNYPQNYFANIEC 835
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1560 LSCQRCWNTSSTEDLPDCHIHMDICWKX-----VTNDVCSEPLRGHRTYIEC 1606

## RESULT 13

S78040  
f:bulin, splice form C precursor - mouse  
N:Alternate names: basement-membrane protein BM-90  
C:Species: Mus musculus (house mouse)  
C:Date: 24-Jul-1998 #sequence revision 24-Jul-1998 #text\_change 02-Aug-2002  
C:Accession: S78040; S78560; S36440  
R:Pan, T.C.; Kluge, M.; Zhang, R.Z.; Mayer, U.; Timp, R.; Chu, M.L.  
Eur. J. Biochem. 215, 733-740, 1993  
A:Title: Sequence of extracellular mouse protein BM-90/fibulin and its calcium-depende  
A:Reference number: S34968; XUID:93358897; PMID:8354280  
A:Accession: S78340  
A:Molecule type: mRNA  
A:Residues: 1-685 <PAN>  
A:Cross-references: EMBL:X70854  
R:Pan, T.C.; Kluge, M.; Zhang, R.Z.; Mayer, U.; Timp, R.; Chu, M.L.  
submitted to the EMBL Data Library, January 1993  
A:Description: Sequence of extracellular mouse protein BM-90/fibulin and its calcium-d  
A:Reference number: S36440  
A:Accession: S78560  
A:Molecule type: mRNA  
A:Residues: 1-39, 'P', 41-685 <CHU>  
A:Cross-references: EMBL:X70854  
C:Genetics:  
A:Introns: 568/3  
C:Superfamily: fibulin-1; EGF homology  
C:Keywords: alternative splicing; basement membrane; calcium binding; extracellular ma  
F:1-29/Domain: signal sequence #status predicted <SIG>  
F:30-695/Product: fibulin, splice form C #status predicted <MAT>  
F:98,537,541/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.1%; Score 567; DB 2; Length 685;

Best Local Similarity 27.7%; Pred. No. 5.9e-27;

Matches 159; Conservative 74; Mismatches 159; Indels 182; Gaps 31;

```
QY 44 EDVDECAQGLDDCHADALCONTPYSKCKPGVQGEGRQ--CEDIDCEGNELNCGC 98
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 217 EDINECTGSHNCRGLGESCINTVGSFRQORDSCGTYELTDNNCKDIDECETGHI-NC 275

QY 99 VHD--C-LNIPGNRYC---TCPDGFMLAHDGHNCLDVDECLNNGGCO--HTCVNVMGSE 150
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 276 PPDFIQLNTLSGRCPKLOCKSGFTDQALG-NCIDNECLSLAPCPVQGTQINTEGSY 334

QY 151 EC-----CKEGEFLSDNQHCT--HRSEGLSCMKNDH-----GCSHICKEAPRGSA 197
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 335 TCQKNVNPNGRGY-----HLNEEGTRCVDVDECAPPAEPCKGHCLNSP-GSFR 383

QY 198 CECRPGFELAKNORDCILTCNHNCGGQCHSDDTADGPESSCHPQYKMHDTGRSCC-LE 257
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 384 CECRAGFYDGIQRTCV-----DINECQRYP-----GRJC----- 423

QY 258 TVLEVTESNTTSVVDGSKRVKRELLMETCAVNNGGCDRTCKDTSTGVHSCVPVGTFLQD 317
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 414 -----GKH-----CENTPGSFHCSCSAGFLSVD 437

QY 318 GYTKCKDIDECQTRNGGCDHFCNKIVSGFDGCKKGFLL--TDEKSCQDVDECSL-----DRT 373
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 438 GRSCEDVNEC--LNSPSCQECANVGSYQYCRRGYQLSDVDGVTCEIDICALPTGGHI 495

QY 374 CDHSC-NHRCPTFACAC-NRGYTYL-QFTHCGDTNEC--SINNGGCGQCVNTVGSYEC-- 427
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 496 CSYEC-NIPGSFQCCSPSSGYRLAPNRCNQDIDECVTGIHNCIS-NETCFN-CGSRCLUS 555

QY 428 -CQHPGVK-----JHNNKCKCEVKGLL-----PTSVS-PRVSLHCKSGSGG 467
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 556 FCEPEYRSADTRCARLPCHEN-GSCPRLPLRITYHLSPFTNIQVPAVFMGSSAV 514

QY 468 DCGFLRCHSHLSSDVTITRTSVTFKLNKGCCLKNAELFPEGLRPAALPEKHSYKESF 527
```

```

Db      615 PG-----DSMQLAITAGNEGFFTRKV-----SHHSGV----- 643
QY      528 RYVNLTCSSGKQVGPAGRPSTPKEMFIVTEFEL 561
Db      644 --VALT---KPIP-----EPRDLTLTVKMDL 664

RESULT 14
C36346
fibulin-1 precursor, splice form C - human
N:Alternate names: fibulin C
N:Contains: fibulin 1 splice form A; fibulin 1 splice form C
C:Species: Homo sapiens (man)
C>Date: 19-Apr-1991 #sequence_revision 19-Apr-1991 #text_change 02-Aug-2002
C:Accession: C36346; A36346; A32826
R:Argaves, W.S.; Tran, H.; Burgess, W.H.; Dickerson, K.
J. Cell Biol. 111, 3155-3164, 1990
A:Title: Fibulin is an extracellular matrix and plasma glycoprotein with repeated domain
A:Reference number: A36346; MUID:91100426; PMID:2265669
A:Accession: C36346
A:Molecule type: mRNA
A:Residues: 1-683 <ARG>
A:Cross-references: GB:X53743; NID:g31418; PIDN:CAA37772.1; PID:g31419
A:Accession: A36346
A:Molecule type: mRNA
A:Residues: 1-566 <AR2>
A:Cross-references: GB:X53741; NID:g31414; PIDN:CAA37770.1; PID:g31415
R:Argaves, W.S.; Dickerson, K.; Burgess, W.H.; Ruoslahti, E.
Cell 58, 623-629, 1989
A:Title: Fibulin, a novel protein that interacts with the fibronectin receptor beta-sub
A:Reference number: A32826; MUID:89354537; PMID:2527614
A:Accession: A32826
A:Molecule type: protein
A:Residues: 30-35, SX, 38-40, 'SH', 43-44 <AR3>
C:Genetics:
A:Gene: GDB:FBLN1; FBLN
A:Cross-references: GDB:278285; OMIM:135820
A:Map position: 22q13.3-22q13.3
C:Superfamily: fibulin-1; EGF homology
C:Keywords: alternative splicing; glycoprotein
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-683/Product: fibulin 1 splice form C #status predicted <MAT>
F:180-214/Domain: EGF homology <EGF>
F:485-523/Domain: EGF homology <EGF>
F:98,535,539/Binding site: carbohydrate (Asn) (covalent: #status predicted

Query Match      10.1%; Score 565; DB 2; Length 683;
Best Local Similarity 27.9%; Pred. No. 7.6e-27;
Matches 160; Conservative 77; Mismatches 155; Indels 192; Gaps 31;

QY      44 EDVDSACGLDDCHADALCONFTSYKC-----SKPGYQ-GEGRQCEIDECGHELAGGC 98
Db      215 EDVNEGIGTGHSCRLGESCINTVSGFRQORSSCGTGHELTEDNSCKDIDECESGIIH-NC 273
QY      99 VHD--CLNIPGNYRC-----TCDFGFMALDHGNCILDVDECIENNGGCQ--HTCVNVMGSSY 150
Db      274 LPDFICQNTLGSFRCPKPLQCKSGFIQDALG-NCIDINECLISAFPCFIGHTC-NTGESSY 332
QY      151 EC-----CKKEGFFLSDNQHTCIHRSEGLSCMNKDH-----GSHICKAPRGSAVA 197
Db      333 TCOKNVNPGCGY-----HUNEETRCVDVDECAPPAPCGKGRGVNPF-GSFR 391
QY      198 CECRPGFELAKNQRCILTCNHNNGGCQHSDDTAJGPECSCHPQYKXHTDGRSCLERED 257
Db      382 CEKGTGYFDGIRMCV-----DVNECQRYP-----GRLC----- 411
QY      258 TVLEVTESNTTSVPDGRKVRKRLMETCAVNNGCCDR'CKDITSTGVHSCSPVGTQLQD 317
Db      * 412 -----GKH-----CENTLGSYLSCSVGFR-LSDV 435
QY      318 GKTCKDIDECQTRNGCDHFCKNIIVGSPDCGCKKGFULL-TDEKSCQDVDECSL---DRT 373

```

```

Db      436 GRSCSDINECS--SSPCGECANVYGSYCCYCRGYCLSDVGVTCEDIDECALPTGGHI 493
QY      374 CMHSCINHPGTACAC-NRGYTLY-GFTHGDTNRC--SINNGCCQCVQVNTVGSYEC-- 427
Db      494 CSYRCINIFGSPQSPSSGYRLAPNGRNCQD'DECV'GIHNCN:NETCFKIQGAFRCJA 553
QY      428 -OCHPGYK-----LHWNKKDCVEVKLL-----PTSV-SPEVSLHCKCKSGG 467
Db      554 FPCPENYRRAATRCERLPCHEN-RECSKLP:R'YYH:LSFPTN:QAFVVPKXGPPSSAV 612
QY      468 DGCFLRCHSGIHLSDEVITIR'SVTFKLNKCKSLKNAELFEGUPPALPEKHSGVKSEF 527
Db      613 PG-----DSMQLAITGNEGFFTRKV-----SPHSGV----- 641
QY      528 RYVNLTCSSGKQVGPAGRPSTPKEMFIVTEFEL 561
Db      642 --VALT---KPIP-----EPRDLTLTVKMDL 662

RESULT 15
A55184
fibulin-2 precursor - human
N:Alternate names: protein DKFP586A1519.1
C:Species: Homo sapiens (man)
C>Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 08-Sep-2002
C:Accession: A55184; T08744
R:Zhang, R.Z.; Ban, T.C.; Zhang, Z.Y.; Mattei, M.G.; Timp, R.; Chu, M.L.
Genomics 22, 425-430, 1994
A:Title: Fibulin-2 (FBLN2): human cDNA sequence, mRNA expression, and mapping of the g
A:Reference number: A55184; MUID:95:04855; PMID:7806230
A:Accession: A55184
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1184 <ZHA>
A:Cross-references: GB:X82494; NID:g575232; PIDN:CAA57876.1; PID:g575233
R:Mambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, March 1995
A:Reference number: Z16471
A:Accession: T08744
A:Molecule type: mRNA
A:Residues: 656-719, QDECLMGARHDCRRORFCVNTLGSFYCVNHTVLCADGY-LNAHRKCVQ', 720-853, 'T', 855-1
A:Cross-references: EMBL:AJ050035
A:Experimental source: adult uterus; clone DKFP586A1519
C:Genetics:
A:Gene: GDB:FBLN2
A:Cross-references: GDB:293037; OMIM:135821
A:Map position: 3p25-3p24
A:Superfamily: fibulin-2; EGF homology; von Willebrand factor type C repeat homology
C:Keywords: alternative splicing; extracellular matrix
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-1184/Product: fibulin-2 protein #status predicted <MAT>
F:955-941/Domain: EGF homology <EGF>

Query Match      10.0%; Score 560; DB 2; Length 1184;
Best Local Similarity 29.1%; Pred. No. 2.7e-26;
Matches 141; Conservative 68; Mismatches 155; Indels 138; Gaps 25;

QY      44 EDVDECAQGLDDCHADALCONFTSYKC-----SKPGYQGEGRQCEIDECG--NELNG 96
Db      718 EC'NECVTLTHTCRSGEHCVNT'GSFHCYKALTCPGVALKDGECEDEECAMGTH'CCP 777
QY      97 GCVHDCNLNIPGNYRC-----TCDFGFMALDHGNCILDVDECIENNGGCQ--HTCVNVMGSSY 150
Db      778 GFL--QNTKGSFYCOARQFCNCGFLQDPBG-NCVDINECTLSLSEPCRPGFSCINTVGSY 834
QY      151 EC-----CKKEGFFLSDNQHTCIHRSEGLSCMNKDHGC--SHICKAPRGSAVACECRPG 203
Db      835 TCORNPICARGYHASDDGAKCVNE-----CETGVHRCGEGQVCHNLP-GSYRCDCKAG 889
QY      204 FSLAKNQRCILTCNHNNGGCQHSDDTAJGPECSCHPQYKXHTDGRSCLEREDTVLEVT 263
Db      890 F-----ORDAF-----GRGCLD----- 901

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2003, 11:17:23 ; Search time 26 seconds  
(without alignments)  
1806.911 Million cell updates/sec

Title: US-09-747-371-2

Perfect score: 5605

Sequence: 1 MGVAGNRNPGAAWALLLL.....PRSFILLRSKYSRFRPYX 999

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	727	13.0	2911	1 FBN2_HUMAN	P35556 homo sapien
2	724	12.9	2907	1 FBN2_MOUSE	Q61555 mus musculus
3	692	12.3	956	1 MTN2_MOUSE	O08746 mus musculus
4	691	12.3	2871	1 FBNI_BOVIN	P98133 bos taurus
5	691	12.3	2871	1 FBNI_HUMAN	P35555 homo sapien
6	690	12.3	2871	1 FBNI_MOUSE	O61554 mus musculus
7	684.5	12.2	956	1 MTN2_HUMAN	O00339 homo sapien
8	678	12.1	2871	1 FBNI_PIG	Q95V36 sus scrofa
9	624.5	11.1	1394	1 LTBS_HUMAN	P22064 homo sapien
10	624.5	11.1	1595	1 LTBL_HUMAN	Q14766 homo sapien
11	601	10.7	1389	1 LTBL_MOUSE	Q8CG18 mus musculus
12	601	10.7	1713	1 LTBL_MOUSE	Q8CG19 mus musculus
13	592	10.6	1121	1 LTBI_RAT	Q00918 rattus norv
14	577.5	10.3	1221	1 FB2_MOUSE	P37889 mus musculus
15	567	10.1	703	1 FB1_HUMAN	P23142 homo sapien
16	562	10.0	704	1 FB1_CHICK	Q73775 gallus gal
17	560	10.0	1184	1 FB2_HUMAN	P98095 homo sapien
18	558	10.0	705	1 FB1_MOUSE	O08879 mus musculus
19	529.5	9.4	798	1 FB1_CAEEL	O77469 caenorhabdi
20	526.5	9.4	2321	1 NTC3_HUMAN	Q9UM47 homo sapien
21	524.5	9.4	598	1 FB1_CERAE	Q8MJJ9 cercopithe
22	524.5	9.4	2318	1 NTC3_MOUSE	Q61982 mus musculus
23	516.5	9.2	2319	1 NTC3_RAT	Q9A172 rattus norv
24	509.5	9.1	2703	1 NOTC_DROME	P07207 drosophila
25	501.5	8.9	2524	1 NOTC_XENLA	P21783 xenopus lae
26	500.5	8.9	681	1 FB1_BRARE	Q42182 brachydanio
27	500	8.9	2470	1 NTC2_MOUSE	C35516 mus musculus
28	489.5	8.7	1964	1 NTC4_MOUSE	P31695 mus musculus
29	489	8.7	2531	1 NTC1_MOUSE	Q01705 mus musculus
30	489	8.7	2531	1 NTC1_RAT	Q07008 rattus norv
31	485	8.7	2471	1 NTC2_HUMAN	Q04721 homo sapien
32	484.5	8.6	2556	1 NTC1_HUMAN	P46531 homo sapien
33	484	8.6	2437	1 NTC1_BRARE	P46530 brachydanio

RESULT 1  
FBN2\_HUMAN 34 479 8.5 2471 1 NTC2\_RAT Q9QW30 rattus norv  
ID FBN2\_HUMAN 35 467 8.3 2003 1 NTC4\_HUMAN Q99466 homo sapien  
AC P35556; 36 466 8.3 1064 1 FBPI\_STRPU P10079 strongyloce  
DT 01-JUN-1994 (Rel. 29, Created) 37 464.5 8.3 3767 1 MUA3\_CAEEL P34576 caenorhabdi  
DT 01-FEB-1996 (Rel. 33, Last sequence update) 38 443.5 7.9 1429 1 L112\_CAEEL P14595 caenorhabdi  
DT 28-FEB-2003 (Rel. 41, Last annotation update) 39 438 7.8 1217 1 EGF\_MOUSE P01132 mus musculus  
DE Fibrillin 2 precursor. 40 434 7.7 448 1 FBUS\_RAT Q9WV08 rattus norv  
GN FBN2. 41 431.5 7.7 1133 1 EGF\_RAT P07522 rattus norv  
OS Homo sapiens (Human). 42 429 7.7 1238 1 JAG2\_HUMAN Q9Y219 homo sapien  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; 43 426.5 7.6 448 1 FBUS\_HUMAN Q9UBX5 homo sapien  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 44 426.5 7.6 448 1 FBUS\_MOUSE Q9WV09 mus musculus  
CX NCBI\_TaxID=9606; 45 421.5 7.5 1218 1 JAG1\_HUMAN P78504 homo sapien  
RN 1.  
SEQUENCE FROM N.A.  
MEDLINE=94165150; PubMed=5120105;  
Zhang H., Apfelroch S.D., Hu W., Davis E.C., Sanguinetti C.,  
Bonadio J., Yecham R.P., Ramirez F.,  
RT "Structure and expression of fibrillin-2, a novel microfibrillar  
RT component preferentially located in elastic matrices.";  
J. Cell Biol. 124:855-863(1994).  
(2)  
SEQUENCE OF 752-1505 FROM N.A.  
MEDLINE=91304567; PubMed=552206;  
Lee B., Godfrey M., Vitale E., Hori H., Mattei M.-G., Sarfarazi M.,  
Raipouras P., Ramirez F., Hollister D.;  
RT "Linkage of Marfan syndrome and a phenotypically related disorder to  
RT two different fibrillin genes.";  
Nature 352:330-334(1991).  
[3]  
VARIANTS CCA TYR-1252 AND SER-1433, AND VARIANT ILE-964.  
MEDLINE=96083599; PubMed=7493032;  
Putnam E.A., Zhang H., Ramirez F., Milewicz D.M.;  
RT "Fibrillin-2 (FBN2) mutations result in the Marfan-like disorder,  
RT congenital contractural arachnoidactyly.";  
Nat. Genet. 11:456-458(1995).  
[4]  
VARIANTS CCA HIS-1114.  
MEDLINE=98407789; PubMed=9737771;  
Babcock D., Gasner C., Francke U., Maslen C.;  
RT "A single mutation that results in an asp-to-his substitution and  
RT partial exon skipping in a family with congenital contractural  
RT arachnoidactyly.";  
Hum. Genet. 103:22-28(1999).  
[5]  
VARIANTS CCA PHE-1141 AND TRP-1252.  
MEDLINE=20259236; PubMed=0797416;  
Beilen S., Zhou G., Wang M., Der Kaloustian V.M., Pagon R.A.,  
Godfrey M.;  
RT "Two novel fibrillin-2 mutations in congenital contractural  
RT arachnoidactyly.";  
Am. J. Med. Genet. 92:7-12(2000).  
CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS  
CC THAT BINDS CALCIUM. FIBRILLIN-2-CONTAINING MICROFIBRILS REGULATE  
CC THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.  
CC -!- DISEASE: DEFECTS IN FBN2 ARE THE CAUSE OF CONGENITAL CONTRACTURAL

CC ARACHNOIDACTYLY (CCA) (ALSO KNOWN AS BEALS SYNDROME). CCA IS  
 CC PHENOTYPICALLY SIMILAR TO MARFAN SYNDROME, BUT DOES NOT EFFECT THE  
 CC AORTA AND THE EYES.  
 CC -!- SIMILARITY: Contains 47 EGF-like domains.  
 CC -!- SIMILARITY: Contains 7 TGF-beta binding protein (TGFBP) domains.  
 CC -!- DATABASE: NAME:Elastic Fiber Homepage; NOTE=fibrillin 2 page;  
 CC WWW="http://ef.wustl.edu/genes/FBN2.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U03272; AAA18950.1; -;  
 CC EMBL; X62009; -; NOT ANNOTATED\_CDS.  
 CC PIR; A54105; A54105.  
 CC HSSP; P35555; 1EMN.  
 CC Genew; HGNC:1604; FBN2.  
 CC MW; 121053; -;  
 CC GO; GO:0005578; C:extracellular matrix; TAS.  
 CC GO; GO:0005201; P:extracellular matrix structural constituent; TAS.  
 CC GO; GO:0007345; P:embryogenesis and morphogenesis; TAS.  
 CC GO; GO:0007397; P:histogenesis and organogenesis; TAS.  
 CC InterPro; IPR000152; ASX\_hydroxyl.  
 CC InterPro; IPR001881; EGF\_Ca.  
 CC InterPro; IPR001438; EGF-II.  
 CC InterPro; IPR006209; EGF-like.  
 CC InterPro; IPR002212; Fibrill-assoc.  
 CC Pfam; PF00308; EGF; 45.  
 CC Pfam; PF00683; TB; 9.  
 CC PRINTS; PR00010; EGFLOOD.  
 CC SMART; SM00179; EGF\_CA; 43.  
 CC PROSITE; PS00010; ASX\_HYDROXYL; 43.  
 CC PROSITE; PS00022; EGF 1; 2.  
 CC PROSITE; PS01186; EGF 2; 37.  
 CC PROSITE; PS01187; EGF\_CA; 42.  
 CC Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;  
 KW Repeat; Signal; Multigene family; Disease mutation; Polymorphism.  
 FT SIGNAL 1 28  
 FT CHAIN 29 2911  
 FT FIBRILLIN 2.  
 FT- DOMAIN 11 142 EGF-LIKE 1.  
 FT- DOMAIN 145 176 EGF-LIKE 2.  
 FT- DOMAIN 176 207 EGF-LIKE 3.  
 FT- DOMAIN 275 316 EGF-LIKE 4.  
 FT- DOMAIN 317 358 EGF-LIKE 5.  
 FT- REPEAT 359 425 TGFBP 1.  
 FT- DOMAIN 493 533 EGF-LIKE 6.  
 FT- DOMAIN 534 573 EGF-LIKE 7.  
 FT- DOMAIN 574 615 EGF-LIKE 8.  
 FT- DOMAIN 616 656 EGF-LIKE 9.  
 FT- DOMAIN 657 697 EGF-LIKE 10.  
 FT- REPEAT 698 766 TGFBP 2.  
 FT- DOMAIN 767 808 EGF-LIKE 11.  
 FT- DOMAIN 809 850 EGF-LIKE 12.  
 FT- DOMAIN 851 890 EGF-LIKE 13.  
 FT- DOMAIN 894 925 EGF-LIKE 14.  
 FT- REPEAT 926 1071 TGFBP 3.  
 FT- DOMAIN 1072 1113 EGF-LIKE 15.  
 FT- DOMAIN 1114 1156 EGF-LIKE 16.  
 FT- DOMAIN 1157 1198 EGF-LIKE 17.  
 FT- DOMAIN 1199 1240 EGF-LIKE 18.  
 FT- DOMAIN 1241 1281 EGF-LIKE 19.  
 FT- DOMAIN 1282 1323 EGF-LIKE 20.  
 FT- DOMAIN 1324 1365 EGF-LIKE 21.  
 FT- DOMAIN 1366 1406 EGF-LIKE 22.  
 FT- DOMAIN 1407 1447 EGF-LIKE 23.  
 FT- DOMAIN 1448 1489 EGF-LIKE 24.  
 FT- DOMAIN 1490 1530 EGF-LIKE 25.  
 FT- DOMAIN 1531 1571 EGF-LIKE 26.  
 FT- REPEAT 1572 1648 TGFBP 4.  
 FT- DOMAIN 1649 1690  
 FT- DOMAIN 1691 1732  
 FT- REPEAT 1733 1806  
 FT- DOMAIN 1807 1848  
 FT- DOMAIN 1849 1890  
 FT- DOMAIN 1891 1932  
 FT- DOMAIN 1933 1974  
 FT- DOMAIN 1975 2016  
 FT- DOMAIN 2017 2058  
 FT- DOMAIN 2059 2099  
 FT- REPEAT 2100 2141  
 FT- DOMAIN 2142 2183  
 FT- DOMAIN 2184 2225  
 FT- DOMAIN 2226 2267  
 FT- DOMAIN 2268 2309  
 FT- DOMAIN 2310 2351  
 FT- DOMAIN 2352 2393  
 FT- DOMAIN 2394 2435  
 FT- REPEAT 2436 2477  
 FT- DOMAIN 2478 2519  
 FT- DOMAIN 2520 2561  
 FT- DOMAIN 2562 2603  
 FT- DOMAIN 2604 2645  
 FT- DOMAIN 2646 2687  
 FT- DISULFID 115 124  
 FT- DISULFID 125 134  
 FT- DISULFID 135 144  
 FT- DISULFID 145 154  
 FT- DISULFID 155 164  
 FT- DISULFID 165 174  
 FT- DISULFID 175 184  
 FT- DISULFID 185 194  
 FT- DISULFID 195 204  
 FT- DISULFID 205 214  
 FT- DISULFID 215 224  
 FT- DISULFID 225 234  
 FT- DISULFID 235 244  
 FT- DISULFID 245 254  
 FT- DISULFID 255 264  
 FT- DISULFID 265 274  
 FT- DISULFID 275 284  
 FT- DISULFID 285 294  
 FT- DISULFID 295 304  
 FT- DISULFID 305 314  
 FT- DISULFID 315 324  
 FT- DISULFID 325 334  
 FT- DISULFID 335 344  
 FT- DISULFID 345 354  
 FT- DISULFID 355 364  
 FT- DISULFID 365 374  
 FT- DISULFID 375 384  
 FT- DISULFID 385 394  
 FT- DISULFID 395 404  
 FT- DISULFID 405 414  
 FT- DISULFID 415 424  
 FT- DISULFID 425 434  
 FT- DISULFID 435 444  
 FT- DISULFID 445 454  
 FT- DISULFID 455 464  
 FT- DISULFID 465 474  
 FT- DISULFID 475 484  
 FT- DISULFID 485 494  
 FT- DISULFID 495 504  
 FT- DISULFID 505 514  
 FT- DISULFID 515 524  
 FT- DISULFID 525 534  
 FT- DISULFID 535 544  
 FT- DISULFID 545 554  
 FT- DISULFID 555 564  
 FT- DISULFID 565 574  
 FT- DISULFID 575 584  
 FT- DISULFID 585 594  
 FT- DISULFID 595 604  
 FT- DISULFID 605 614  
 FT- DISULFID 615 624  
 FT- DISULFID 625 634  
 FT- DISULFID 635 644  
 FT- DISULFID 645 654  
 FT- DISULFID 655 664  
 FT- DISULFID 665 674  
 FT- DISULFID 675 684  
 FT- DISULFID 685 694  
 FT- DISULFID 695 704  
 FT- DISULFID 705 714  
 FT- DISULFID 715 724  
 FT- DISULFID 725 734  
 FT- DISULFID 735 744  
 FT- DISULFID 745 754  
 FT- DISULFID 755 764  
 FT- DISULFID 765 774  
 FT- DISULFID 775 784  
 FT- DISULFID 785 794  
 FT- DISULFID 795 804  
 FT- DISULFID 805 814  
 FT- DISULFID 815 824  
 FT- DISULFID 825 834  
 FT- DISULFID 835 844  
 FT- DISULFID 845 854  
 FT- DISULFID 855 864  
 FT- DISULFID 865 874  
 FT- DISULFID 875 884  
 FT- DISULFID 885 894  
 FT- DISULFID 895 904  
 FT- DISULFID 905 914  
 FT- DISULFID 915 924  
 FT- DISULFID 925 934  
 FT- DISULFID 935 944  
 FT- DISULFID 945 954  
 FT- DISULFID 955 964  
 FT- DISULFID 965 974  
 FT- DISULFID 975 984  
 FT- DISULFID 985 994  
 FT- DISULFID 995 1004  
 FT- DISULFID 1005 1014  
 FT- DISULFID 1015 1024  
 FT- DISULFID 1025 1034  
 FT- DISULFID 1035 1044  
 FT- DISULFID 1045 1054  
 FT- DISULFID 1055 1064  
 FT- DISULFID 1065 1074  
 FT- DISULFID 1075 1084  
 FT- DISULFID 1085 1094  
 FT- DISULFID 1095 1104  
 FT- DISULFID 1105 1114  
 FT- DISULFID 1115 1124  
 FT- DISULFID 1125 1134  
 FT- DISULFID 1135 1144  
 FT- DISULFID 1145 1154  
 FT- DISULFID 1155 1164  
 FT- DISULFID 1165 1174  
 FT- DISULFID 1175 1184  
 FT- DISULFID 1185 1194  
 FT- DISULFID 1195 1204  
 FT- DISULFID 1205 1214  
 FT- DISULFID 1215 1224  
 FT- DISULFID 1225 1234  
 FT- DISULFID 1235 1244  
 FT- DISULFID 1245 1254  
 FT- DISULFID 1255 1264  
 FT- DISULFID 1265 1274  
 FT- DISULFID 1275 1284  
 FT- DISULFID 1285 1294  
 FT- DISULFID 1295 1304  
 FT- DISULFID 1305 1314  
 FT- DISULFID 1315 1324  
 FT- DISULFID 1325 1334  
 FT- DISULFID 1335 1344  
 FT- DISULFID 1345 1354  
 FT- DISULFID 1355 1364  
 FT- DISULFID 1365 1374  
 FT- DISULFID 1375 1384  
 FT- DISULFID 1385 1394  
 FT- DISULFID 1395 1404  
 FT- DISULFID 1405 1414  
 FT- DISULFID 1415 1424  
 FT- DISULFID 1425 1434  
 FT- DISULFID 1435 1444  
 FT- DISULFID 1445 1454  
 FT- DISULFID 1455 1464  
 FT- DISULFID 1465 1474  
 FT- DISULFID 1475 1484  
 FT- DISULFID 1485 1494  
 FT- DISULFID 1495 1504  
 FT- DISULFID 1505 1514  
 FT- DISULFID 1515 1524  
 FT- DISULFID 1525 1534  
 FT- DISULFID 1535 1544  
 FT- DISULFID 1545 1554  
 FT- DISULFID 1555 1564  
 FT- DISULFID 1565 1574  
 FT- DISULFID 1575 1584  
 FT- DISULFID 1585 1594  
 FT- DISULFID 1595 1604  
 FT- DISULFID 1605 1614  
 FT- DISULFID 1615 1624  
 FT- DISULFID 1625 1634  
 FT- DISULFID 1635 1644  
 FT- DISULFID 1645 1654  
 FT- DISULFID 1655 1664  
 FT- DISULFID 1665 1674  
 FT- DISULFID 1675 1684  
 FT- DISULFID 1685 1694  
 FT- DISULFID 1695 1704  
 FT- DISULFID 1705 1714  
 FT- DISULFID 1715 1724  
 FT- DISULFID 1725 1734  
 FT- DISULFID 1735 1744  
 FT- DISULFID 1745 1754  
 FT- DISULFID 1755 1764  
 FT- DISULFID 1765 1774  
 FT- DISULFID 1775 1784  
 FT- DISULFID 1785 1794  
 FT- DISULFID 1795 1804  
 FT- DISULFID 1805 1814  
 FT- DISULFID 1815 1824  
 FT- DISULFID 1825 1834  
 FT- DISULFID 1835 1844  
 FT- DISULFID 1845 1854  
 FT- DISULFID 1855 1864  
 FT- DISULFID 1865 1874  
 FT- DISULFID 1875 1884  
 FT- DISULFID 1885 1894  
 FT- DISULFID 1895 1904  
 FT- DISULFID 1905 1914  
 FT- DISULFID 1915 1924  
 FT- DISULFID 1925 1934  
 FT- DISULFID 1935 1944  
 FT- DISULFID 1945 1954  
 FT- DISULFID 1955 1964  
 FT- DISULFID 1965 1974  
 FT- DISULFID 1975 1984  
 FT- DISULFID 1985 1994  
 FT- DISULFID 1995 2004  
 FT- DISULFID 2005 2014  
 FT- DISULFID 2015 2024  
 FT- DISULFID 2025 2034  
 FT- DISULFID 2035 2044  
 FT- DISULFID 2045 2054  
 FT- DISULFID 2055 2064  
 FT- DISULFID 2065 2074  
 FT- DISULFID 2075 2084  
 FT- DISULFID 2085 2094  
 FT- DISULFID 2095 2104  
 FT- DISULFID 2105 2114  
 FT- DISULFID 2115 2124  
 FT- DISULFID 2125 2134  
 FT- DISULFID 2135 2144  
 FT- DISULFID 2145 2154  
 FT- DISULFID 2155 2164  
 FT- DISULFID 2165 2174  
 FT- DISULFID 2175 2184  
 FT- DISULFID 2185 2194  
 FT- DISULFID 2195 2204  
 FT- DISULFID 2205 2214  
 FT- DISULFID 2215 2224  
 FT- DISULFID 2225 2234  
 FT- DISULFID 2235 2244  
 FT- DISULFID 2245 2254  
 FT- DISULFID 2255 2264  
 FT- DISULFID 2265 2274  
 FT- DISULFID 2275 2284  
 FT- DISULFID 2285 2294  
 FT- DISULFID 2295 2304  
 FT- DISULFID 2305 2314  
 FT- DISULFID 2315 2324  
 FT- DISULFID 2325 2334  
 FT- DISULFID 2335 2344  
 FT- DISULFID 2345 2354  
 FT- DISULFID 2355 2364  
 FT- DISULFID 2365 2374  
 FT- DISULFID 2375 2384  
 FT- DISULFID 2385 2394  
 FT- DISULFID 2395 2404  
 FT- DISULFID 2405 2414  
 FT- DISULFID 2415 2424  
 FT- DISULFID 2425 2434  
 FT- DISULFID 2435 2444  
 FT- DISULFID 2445 2454  
 FT- DISULFID 2455 2464  
 FT- DISULFID 2465 2474  
 FT- DISULFID 2475 2484  
 FT- DISULFID 2485 2494  
 FT- DISULFID 2495 2504  
 FT- DISULFID 2505 2514  
 FT- DISULFID 2515 2524  
 FT- DISULFID 2525 2534  
 FT- DISULFID 2535 2544  
 FT- DISULFID 2545 2554  
 FT- DISULFID 2555 2564  
 FT- DISULFID 2565 2574  
 FT- DISULFID 2575 2584  
 FT- DISULFID 2585 2594  
 FT- DISULFID 2595 2604  
 FT- DISULFID 2605 2614  
 FT- DISULFID 2615 2624  
 FT- DISULFID 2625 2634  
 FT- DISULFID 2635 2644  
 FT- DISULFID 2645 2654  
 FT- DISULFID 2655 2664  
 FT- DISULFID 2665 2674  
 FT- DISULFID 2675 2684  
 FT- DISULFID 2685 2694  
 FT- DISULFID 2695 2704  
 FT- DISULFID 2705 2714  
 FT- DISULFID 2715 2724  
 FT- DISULFID 2725 2734  
 FT- DISULFID 2735 2744  
 FT- DISULFID 2745 2754  
 FT- DISULFID 2755 2764  
 FT- DISULFID 2765 2774  
 FT- DISULFID 2775 2784  
 FT- DISULFID 2785 2794  
 FT- DISULFID 2795 2804  
 FT- DISULFID 2805 2814  
 FT- DISULFID 2815 2824  
 FT- DISULFID 2825 2834  
 FT- DISULFID 2835 2844  
 FT- DISULFID 2845 2854  
 FT- DISULFID 2855 2864  
 FT- DISULFID 2865 2874  
 FT- DISULFID 2875 2884  
 FT- DISULFID 2885 2894  
 FT- DISULFID 2895 2904  
 FT- DISULFID 2905 2914  
 FT- DISULFID 2915 2924  
 FT- DISULFID 2925 2934  
 FT- DISULFID 2935 2944  
 FT- DISULFID 2945 2954  
 FT- DISULFID 2955 2964  
 FT- DISULFID 2965 2974  
 FT- DISULFID 2975 2984  
 FT- DISULFID 2985 2994  
 FT- DISULFID 2995 3004  
 FT- DISULFID 3005 3014  
 FT- DISULFID 3015 3024  
 FT- DISULFID 3025 3034  
 FT- DISULFID 3035 3044  
 FT- DISULFID 3045 3054  
 FT- DISULFID 3055 3064  
 FT- DISULFID 3065 3074  
 FT- DISULFID 3075 3084  
 FT- DISULFID 3085 3094  
 FT- DISULFID 3095 3104  
 FT- DISULFID 3105 3114  
 FT- DISULFID 3115 3124  
 FT- DISULFID 3125 3134  
 FT- DISULFID 3135 3144  
 FT- DISULFID 3145 3154  
 FT- DISULFID 3155 3164  
 FT- DISULFID 3165 3174  
 FT- DISULFID 3175 3184  
 FT- DISULFID 3185 3194  
 FT- DISULFID 3195 3204  
 FT- DISULFID 3205 3214  
 FT- DISULFID 3215 3224  
 FT- DISULFID 3225 3234  
 FT- DISULFID 3235 3244  
 FT- DISULFID 3245 3254  
 FT- DISULFID 3255 3264  
 FT- DISULFID 3265 3274  
 FT- DISULFID 3275 3284  
 FT- DISULFID 3285 3294  
 FT- DISULFID 3295 3304  
 FT- DISULFID 3305 3314  
 FT- DISULFID 3315 3324  
 FT- DISULFID 3325 3334  
 FT- DISULFID 3335 3344  
 FT- DISULFID 3345 3354  
 FT- DISULFID 3355 3364  
 FT- DISULFID 3365 3374  
 FT- DISULFID 3375 3384  
 FT- DISULFID 3385 3394  
 FT- DISULFID 3395 3404  
 FT- DISULFID 3405 3414  
 FT- DISULFID 3415 3424  
 FT- DISULFID 3425 3434  
 FT- DISULFID 3435 3444  
 FT- DISULFID 3445 3454  
 FT- DISULFID 3455 3464  
 FT- DISULFID 3465 3474  
 FT- DISULFID 3475 3484  
 FT- DISULFID 3485 3494  
 FT- DISULFID 3495 3504  
 FT- DISULFID 3505 3514  
 FT- DISULFID 3515 3524  
 FT- DISULFID 3525 3534  
 FT- DISULFID 3535 3544  
 FT- DISULFID 3545 3554  
 FT- DISULFID 3555 3564  
 FT- DISULFID 3565 3574  
 FT- DISULFID 3575 3584  
 FT- DISULFID 3585 3594  
 FT- DISULFID 3595 3604  
 FT- DISULFID 3605 3614  
 FT- DISULFID 3615 3624  
 FT- DISULFID 3625 3634  
 FT- DISULFID 3635 3644  
 FT- DISULFID 3645 3654  
 FT- DISULFID 3655 3664  
 FT- DISULFID 3665 3674  
 FT- DISULFID 3675 3684  
 FT- DISULFID 3685 3694  
 FT- DISULFID 3695 3704  
 FT- DISULFID 3705 3714  
 FT- DISULFID 3715 3724  
 FT- DISULFID 3725 3734  
 FT- DISULFID 3735 3744  
 FT- DISULFID 3745 3754  
 FT- DISULFID 3755 3764  
 FT- DISULFID 3765 3774  
 FT- DISULFID 3775 3784  
 FT- DISULFID 3785 3794  
 FT- DISULFID 3795 3804  
 FT- DISULFID 3805 3814  
 FT- DISULFID 3815 3824  
 FT- DISULFID 3825 3834  
 FT- DISULFID 3835 3844  
 FT- DISULFID 3845 3854  
 FT- DISULFID 3855 3864  
 FT- DISULFID 3865 3874  
 FT- DISULFID 3875 3884  
 FT- DISULFID 3885 3894  
 FT- DISULFID 3895 3904  
 FT- DISULFID 3905 3914  
 FT- DISULFID 3915 3924  
 FT- DISULFID 3925 3934  
 FT- DISULFID 3935 3944  
 FT- DISULFID 3945 3954  
 FT- DISULFID 3955 3964  
 FT- DISULFID 3965 3974  
 FT- DISULFID 3975 3984  
 FT- DISULFID 3985 3994  
 FT- DISULFID 3995 4004  
 FT- DISULFID 4005 4014  
 FT- DISULFID 4015 4024  
 FT- DISULFID 4025 4034  
 FT- DISULFID 4035 4044  
 FT- DISULFID 4045 4054  
 FT- DISULFID 4055 4064  
 FT- DISULFID 4065 4074  
 FT- DISULFID 4075 4084  
 FT- DISULFID 4085 4094  
 FT- DISULFID 4095 4104  
 FT- DISULFID 4105 4114  
 FT- DISULFID 4115 4124  
 FT- DISULFID 4125 4134  
 FT- DISULFID 4135 4144  
 FT- DISULFID 4145 4154  
 FT- DISULFID 4155 4164  
 FT- DISULFID 4165 4174  
 FT- DISULFID 4175 4184  
 FT- DISULFID 4185 4194  
 FT- DISULFID 4195 4204  
 FT- DISULFID 4205 4214  
 FT- DISULFID 4215 4224  
 FT- DISULFID 4225 4234  
 FT- DISULFID 4235 4244  
 FT- DISULFID 4245 4254  
 FT- DISULFID 4255 4264  
 FT- DISULFID 4265 4274  
 FT- DISULFID 4275 4284  
 FT- DISULFID 4285 4294  
 FT- DISULFID 4295 4304  
 FT- DISULFID 4305 4314  
 FT- DISULFID 4315 4324  
 FT- DISULFID 4325 4334  
 FT- DISULFID 4335 4344  
 FT- DISULFID 4345 4354  
 FT- DISULFID 4355 4364  
 FT- DISULFID 4365 4374  
 FT- DISULFID 4375 4384  
 FT- DISULFID 4385 4394  
 FT- DISULFID 4395 4404  
 FT- DISULFID 4405 4414  
 FT- DISULFID 4415 4424  
 FT- DISULFID 4425 4434  
 FT- DISULFID 4435 4444  
 FT- DISULFID 4445 4454  
 FT- DISULFID 4455 4464  
 FT- DISULFID 4465 4474  
 FT- DISULFID 4475 4484  
 FT- DISULFID 4485 4494  
 FT- DISULFID 4495 4504  
 FT- DISULFID 4505 4514  
 FT- DISULFID 4515 4524  
 FT- DISULFID 4525 4534  
 FT- DISULFID 4535 4544  
 FT- DISULFID 4545 4554  
 FT- DISULFID 4555 4564  
 FT- DISULFID 4565 4574  
 FT- DISULFID 4575 4584  
 FT- DISULFID 4585 4594  
 FT- DISULFID 4595 4604  
 FT- DISULFID 4605 4614  
 FT- DISULFID 4615 4624  
 FT- DISULFID 4625 4634  
 FT- DISULFID 4635 4644  
 FT- DISULFID 4645 4654  
 FT- DISULFID 4655 4664  
 FT- DISULFID 4665 4674  
 FT- DISULFID 4675 4684  
 FT- DISULFID 4685 4694  
 FT- DISULFID 4695 4704  
 FT- DISULFID 4705 4714  
 FT- DISULFID 4715 4724  
 FT- DISULFID 4725 4734  
 FT- DISULFID 4735 4744  
 FT- DISULFID 4745 4754  
 FT- DISULFID 4755 4764  
 FT- DISULFID 4765 4774  
 FT- DISULFID 4775 4784  
 FT- DISULFID 4785 4794  
 FT- DISULFID 4795 4804  
 FT- DISULFID 4805 4814  
 FT- DISULFID 4815 4824  
 FT- DISULFID 4825 4834  
 FT- DISULFID 4835 4844  
 FT- DISULFID 4845 4854  
 FT- DISULFID 4855 4864  
 FT- DISULFID 4865 4874  
 FT- DISULFID 4875 4884  
 FT- DISULFID 4885 4894  
 FT- DISULFID 4895 4904  
 FT- DISULFID 4905 4914  
 FT- DISULFID 4915 4924  
 FT- DISULFID 4925 4934  
 FT- DISULFID 4935 4944  
 FT- DISULFID 4945 4954  
 FT- DISULFID 4955 4964  
 FT- DISULFID 4965 4974  
 FT- DISULFID 4975 4984  
 FT- DISULFID 4985 4994  
 FT- DISULFID 4995 5004  
 FT- DISULFID 5005 5014  
 FT- DISULFID 5015 5024  
 FT- DISULFID 5025 5034  
 FT- DISULFID 5035 5044  
 FT- DISULFID 5045 5054  
 FT- DISULFID 5055 5064  
 FT- DISULFID 5065 5074  
 FT- DISULFID 5075 5084  
 FT- DISULFID 5085 5094  
 FT- DISULFID 5095 5104  
 FT- DISULFID 5105 5114  
 FT- DISULFID 5115 5124  
 FT- DISULFID 5125 5134  
 FT- DISULFID 5135 5144  
 FT- DISULFID 5145 5154  
 FT- DISULFID 5155 5164  
 FT- DISULFID 5165 5174  
 FT- DISULFID 5175 5184  
 FT- DISULFID 5185 5194  
 FT- DISULFID 5195 5204  
 FT- DISULFID 5205 5214  
 FT- DISULFID 5215 5224  
 FT- DISULFID 5225 5234  
 FT- DISULFID 5235 5244  
 FT- DISULFID 5245 5254  
 FT- DISULFID 5255 5264  
 FT- DISULFID 5265 5274  
 FT- DISULFID 52

Query Match 13.0%; Score 727; DB 1; Length 2911;  
Best Local Similarity 26.5%; Pred. No. 2e-40;  
Matches 255; Conservative 94; Mismatches 301; Indels 312; Gaps 51;

```
QY 45 DVDECAQGGDDCHADALCONTFTSYKSCCKPYOG--EGRCQEDIDECNEELGCGVHDC 102
D 1199 DINECSLS-DNLCRNKGVNMTGTQVCSNPGYQATPDQGTDDDEC-MXMGSGCDTQC 1256
QY 103 LNIPIGYRCTCFDGFMLAHGHNCLDVDECLNN-----GGQHTCVNVMGSEYCCCKEG 157
D 1257 TSESGSEYSCSEGVYALPDGSCADIDEC-ENNEDI CDGG--OCTNIPGEYRCLCYDG 1312
QY 158 FFLSNQHTCIHRSEBGLSCMKDHGCSHICK-----EAPRGVACECPGPEFLANKQDC 213
D 1313 FNASMDMKTCIDVNECDLN-----SNICMFGECENTKGFICHCGLGYSVKVGTTC 1354
QY 214 --LLTCHNGGC--QHSCTDADGPECSCHPQYXMHDTGRSCLERETVLEVTESN-TS 269
D 1365 TDVBECEI GAHNCMDHASCLNIPGFKSCREGM--IGNGIKCIDLDEC-----SNGT- 1415
QY 270 VVDGKRVRRLLMETCAVNGGCCORTCKDTSTGVHSCCPVGTTLQDQKTKCKDIDECCT 329
D 1416 -----HCCSIN-----AQCNTPGSYRCACSEGT--GGFTCSVDVDECAE 1454
QY 330 RNGGCDH-FCKNIVGFDGCKKGFLLTDEKSCQDVDECSLDRTC-DHSCINHPGTAC 387
D 1455 NINLCENGQCLANVPAYRCECEMGFTPASDSRSCODIDECSPQNICVSTCNLPSVFEHC 1514
QY 388 ACNRGYTL-YGTHCGTNECS--IN--NGGQQVCVNTVSGVECOCHPYKHLWNKDC 442
D 1515 ICDDGVELDRTGNCNTDIDECADPINCVG--LCVNTPGRYECNCFDFOLNPTGVGC 1570
QY 443 VEYK-GLLPTSVSPVSLHCGSGGDCGFLRCHSGIHLSSDVTIRTSVTFKLNKGRCS 501
D 1571 VDNRVGNCYLKGR-----GDGS-LSNTEI-----GVGYSRSCC-----CS 1608
QY 502 LKNAELPPGLPALPEKHSVKESPRYNLTCSGKQVGPAGRPSTPKENFTVSEFEL 561
D 1609 LKAWGNPCETCPV-----NSTEYTL-----CPGGEGRPNP-----IITLED 1649
QY 562 ETNQEVTASCD-LSGIVKRTKRLKARLRLKAVHREQFHLQLSGMKLDVAKKPEPVS 620
D 1650 IDECQELPLGCOGNCI-----NOCVS-----CRAGTYDGASERCILCP-----N 1667
QY 621 EREAECGVGGH--AE-----NOCVS-----CRAGTYDGASERCILCP----- 658
D 1668 TFGSFQCECPQYLSYSEDTRICEDICFAHPGVCGEGTCYNTLGNVTCICPEYVQVNG 1727
QY 659 -----NG-TFQNE-----EGQNTCEPCRPQNSG-- 681
D 1728 GHNCMDMRKSFYRSYNGTTCENELPFNVTKRMCOCTYVNGKAGNKPCEPCPTGTADFK 1787
QY 682 -----ALKTPAEMNMBEC----- 694
D 1788 TICGNIPGFTDILHTGKAVDIDECKEIPGICANGVCINQISFRCEBCTGFSYNDLLVC 1847
QY 695 -----GGLCQ-----PGEYSADGAPFQCLCALG-TFQPEAG--RTSCFP-- 730
D 1848 EDIDECNGDNLCQRNADCNISPGSYRCE-----CAAGFKLSPNGACVDNNECLEIPN 1900
QY 731 -CGGGLATKHQATSEFQ-DETRVQCSPGHFYNTTTHRCIR--CPVGTVPQEPGKNCVUS 786
D 1901 VCSHGLCVDLQG--SYQICHNHGFKASQDOTCMQDVDECEHPCGNGTKNTVGSYNCLC 1958
QY 787 CPG-----NTTDFDPSGTNI--TCKNRRCGGLGDF-----TGYESPNYPGVNPANT 833
D 1959 YPQFELTHNNDCLDIDECSSFFQVCNRCNRCNEIGSKFLCONEGYELTPDQKNCIDN- 2017
QY 834 EC 835
D 2018 EC 2019
```

```
RESULT 2
FBN2 MOUSE
ID FBN2 MOUSE STANDARD; PRT; 2907 AA.
AC Q61555; Q63957;
DT 15-DEC-1998 (Rel. 37, Created);
DT 15-DEC-1998 (Rel. 37, Last sequence update);
DT 28-FEB-2003 (Rel. 41, Last annotation update);
DE Fibrillin 2 precursor.
GN FBN2 OR FBN-2.
CS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
PN [?];
SP SEQUENCE FROM N.A.
RX MEDLINE=95263670; PubMed=7744963;
RA Zhang H., Hu W., Ramirez P.;
RT "Developmental expression of fibrillin genes suggests heterogeneity
RT of extracellular microfibrils.";
RJ J. Cell Biol. 129:1165-1176(1995).
RN [?];
RP SEQUENCE OF 210-317 PROX N.A.
RX MEDLINE=94140368; PubMed=8307578;
RA Li X., Pereira L., Zhang H., Sanguinetti C., Ramirez F., Bonadio J.,
RA Francke U.;
RT "Fibrillin genes map to regions of conserved mouse/human synteny on
RT mouse chromosomes 2 and 18.";
RL Genomics 18:667-672(1993).
CC -- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE
CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
CC -- SIMILARITY: Contains 47 EGF-like domains.
CC -- SIMILARITY: Contains 7 TGF-beta binding protein (TGFBR) domains.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
EMBL; L39790; AAA74908.1; -
EMBL; S69359; AAC60685.1; -
PIR; A57278; A57278.
HSP; P35555; IEMN.
MGD; MGI:95490; Fbr2.
GO; GO:0030326; P, limb morphogenesis; IMP.
InterPro; IPR000152; Asx hydroxyl.
InterPro; IPR001861; EGF_Ca.
InterPro; IPR001438; EGF_I.
InterPro; IPR006209; EGF-like.
InterPro; IPR002212; Fibrin-assoc.
Pfam; PF00308; EGF_45.
Pfam; PF00683; TB; 9.
PRINTS; PR00010; EGFBL00D.
SMART; SM00179; EGF_CA; 43.
PROSITE; PS00010; ASX_HYDROXYL; 43.
PROSITE; PS00022; EGF_1; 2.
PROSITE; PS01186; EGF_2; 36.
PROSITE; PS01187; EGF_CA; 43.
KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
Repeat; Signal; Multigene family.
FT SIGNAL 28
FT CHAIN 29 2907 FIBRILLIN 2.
FT DOMAIN 112 142 EGF-LIKE 1.
FT DOMAIN 145 176 EGF-LIKE 2.
FT DOMAIN 176 208 EGF-LIKE 3.
FT DOMAIN 276 317 EGF-LIKE 4.
FT DOMAIN 318 359 EGF-LIKE 5.
FT REPEAT 360 426 TGFSP 1.
FT DOMAIN 487 527 EGF-LIKE 6.
```

```
FT DOMAIN 528 567 EGF-LIKE 7, CALCIUM-BINDING.
FT DOMAIN 568 609 EGF-LIKE 8, CALCIUM-BINDING.
FT DOMAIN 610 650 EGF-LIKE 9, CALCIUM-BINDING.
FT DOMAIN 651 691 EGF-LIKE 10, CALCIUM-BINDING.
FT REPEAT 692 760 TGFBP 2.
FT DOMAIN 761 802 EGF-LIKE 11, CALCIUM-BINDING.
FT DOMAIN 803 844 EGF-LIKE 12, CALCIUM-BINDING.
FT DOMAIN 845 883 EGF-LIKE 13, CALCIUM-BINDING.
FT DOMAIN 948 989 EGF-LIKE 14, CALCIUM-BINDING.
FT REPEAT 990 1065 TGFBP 3.
FT DOMAIN 1066 1107 EGF-LIKE 15, CALCIUM-BINDING.
FT DOMAIN 1108 1150 EGF-LIKE 16, CALCIUM-BINDING.
FT DOMAIN 1151 1192 EGF-LIKE 17, CALCIUM-BINDING.
FT DOMAIN 1193 1234 EGF-LIKE 18, CALCIUM-BINDING.
FT DOMAIN 1235 1275 EGF-LIKE 19, CALCIUM-BINDING.
FT DOMAIN 1276 1317 EGF-LIKE 20, CALCIUM-BINDING.
FT DOMAIN 1318 1359 EGF-LIKE 21, CALCIUM-BINDING.
FT DOMAIN 1360 1400 EGF-LIKE 22, CALCIUM-BINDING.
FT DOMAIN 1401 1441 EGF-LIKE 23, CALCIUM-BINDING.
FT DOMAIN 1442 1483 EGF-LIKE 24, CALCIUM-BINDING.
FT DOMAIN 1484 1524 EGF-LIKE 25, CALCIUM-BINDING.
FT DOMAIN 1525 1565 EGF-LIKE 26, CALCIUM-BINDING.
FT REPEAT 1566 1642 TGFBP 4.
FT DOMAIN 1643 1684 EGF-LIKE 27, CALCIUM-BINDING.
FT DOMAIN 1685 1726 EGF-LIKE 28, CALCIUM-BINDING.
FT REPEAT 1727 1800 TGFBP 5.
FT DOMAIN 1801 1842 EGF-LIKE 29, CALCIUM-BINDING.
FT DOMAIN 1843 1884 EGF-LIKE 30, CALCIUM-BINDING.
FT DOMAIN 1885 1926 EGF-LIKE 31, CALCIUM-BINDING.
FT DOMAIN 1927 1965 EGF-LIKE 32, CALCIUM-BINDING.
FT DOMAIN 1966 2008 EGF-LIKE 33, CALCIUM-BINDING.
FT DOMAIN 2009 2048 EGF-LIKE 34, CALCIUM-BINDING.
FT DOMAIN 2049 2090 EGF-LIKE 35, CALCIUM-BINDING.
FT REPEAT 2091 2163 TGFBP 6.
FT DOMAIN 2164 2205 EGF-LIKE 36, CALCIUM-BINDING.
FT DOMAIN 2206 2245 EGF-LIKE 37, CALCIUM-BINDING.
FT DOMAIN 2246 2286 EGF-LIKE 38, CALCIUM-BINDING.
FT DOMAIN 2287 2330 EGF-LIKE 39, CALCIUM-BINDING.
FT DOMAIN 2331 2372 EGF-LIKE 40, CALCIUM-BINDING.
FT REPEAT 2373 2441 TGFBP 7.
FT DOMAIN 2442 2483 EGF-LIKE 41, CALCIUM-BINDING.
FT DOMAIN 2484 2524 EGF-LIKE 42, CALCIUM-BINDING.
FT DOMAIN 2525 2563 EGF-LIKE 43, CALCIUM-BINDING.
FT DOMAIN 2564 2606 EGF-LIKE 44, CALCIUM-BINDING.
FT DOMAIN 2607 2646 EGF-LIKE 45, CALCIUM-BINDING.
FT DOMAIN 2647 2687 EGF-LIKE 46, CALCIUM-BINDING.
FT DOMAIN 2688 2727 EGF-LIKE 47, CALCIUM-BINDING.
FT DISULFID 115 124 BY SIMILARITY.
FT DISULFID 119 130 BY SIMILARITY.
FT DISULFID 132 141 BY SIMILARITY.
FT DISULFID 149 159 BY SIMILARITY.
FT DISULFID 153 164 BY SIMILARITY.
FT DISULFID 166 175 BY SIMILARITY.
FT DISULFID 180 190 BY SIMILARITY.
FT DISULFID 184 196 BY SIMILARITY.
FT DISULFID 198 207 BY SIMILARITY.
FT DISULFID 280 292 BY SIMILARITY.
FT DISULFID 287 301 BY SIMILARITY.
FT DISULFID 303 316 BY SIMILARITY.
FT DISULFID 322 334 BY SIMILARITY.
FT DISULFID 329 343 BY SIMILARITY.
FT DISULFID 345 358 BY SIMILARITY.
FT DISULFID 498 513 BY SIMILARITY.
FT DISULFID 514 526 BY SIMILARITY.
FT DISULFID 532 542 BY SIMILARITY.
FT DISULFID 537 551 BY SIMILARITY.
FT DISULFID 553 566 BY SIMILARITY.
FT DISULFID 572 584 BY SIMILARITY.
FT DISULFID 579 593 BY SIMILARITY.
FT DISULFID 595 608 BY SIMILARITY.
FT DISULFID 614 625 BY SIMILARITY.
FT DISULFID 620 634 BY SIMILARITY.
```

```
FT DISULFID 636 643 BY SIMILARITY.
FT DISULFID 655 666 BY SIMILARITY.
FT DISULFID 661 675 BY SIMILARITY.
FT DISULFID 677 690 BY SIMILARITY.
FT DISULFID 765 777 BY SIMILARITY.
FT DISULFID 772 786 BY SIMILARITY.
FT DISULFID 788 801 BY SIMILARITY.
FT DISULFID 807 819 BY SIMILARITY.
FT DISULFID 824 828 BY SIMILARITY.
FT DISULFID 832 843 BY SIMILARITY.
FT DISULFID 849 858 BY SIMILARITY.
FT DISULFID 854 868 BY SIMILARITY.
FT DISULFID 870 883 BY SIMILARITY.
FT DISULFID 952 964 BY SIMILARITY.
FT DISULFID 959 973 BY SIMILARITY.
FT DISULFID 975 988 BY SIMILARITY.
FT DISULFID 1070 1082 BY SIMILARITY.
FT DISULFID 1077 1091 BY SIMILARITY.
FT DISULFID 1093 1106 BY SIMILARITY.
FT DISULFID 1112 1124 BY SIMILARITY.
FT DISULFID 1119 1133 BY SIMILARITY.
FT DISULFID 1135 1149 BY SIMILARITY.
FT DISULFID 1155 1167 BY SIMILARITY.
FT DISULFID 1162 1176 BY SIMILARITY.
FT DISULFID 1178 1191 BY SIMILARITY.
FT DISULFID 1197 1209 BY SIMILARITY.
FT DISULFID 1204 1218 BY SIMILARITY.
FT DISULFID 1220 1233 BY SIMILARITY.
FT DISULFID 1239 1250 BY SIMILARITY.
FT DISULFID 1246 1259 BY SIMILARITY.
FT DISULFID 1261 1274 BY SIMILARITY.
FT DISULFID 1280 1292 BY SIMILARITY.
FT DISULFID 1287 1301 BY SIMILARITY.
FT DISULFID 1303 1316 BY SIMILARITY.
FT DISULFID 1322 1334 BY SIMILARITY.
FT DISULFID 1329 1343 BY SIMILARITY.
FT DISULFID 1345 1358 BY SIMILARITY.
FT DISULFID 1364 1377 BY SIMILARITY.
FT DISULFID 1371 1386 BY SIMILARITY.
FT DISULFID 1388 1399 BY SIMILARITY.
FT DISULFID 1405 1418 BY SIMILARITY.
FT DISULFID 1412 1427 BY SIMILARITY.
FT DISULFID 1429 1440 BY SIMILARITY.
FT DISULFID 1446 1458 BY SIMILARITY.
FT DISULFID 1453 1467 BY SIMILARITY.
FT DISULFID 1463 1482 BY SIMILARITY.
FT DISULFID 1469 1499 BY SIMILARITY.
FT DISULFID 1494 1508 BY SIMILARITY.
FT DISULFID 1510 1523 BY SIMILARITY.
FT DISULFID 1529 1542 BY SIMILARITY.
FT DISULFID 1535 1549 BY SIMILARITY.
FT DISULFID 1551 1564 BY SIMILARITY.
FT DISULFID 1647 1653 BY SIMILARITY.
FT DISULFID 1654 1669 BY SIMILARITY.
FT DISULFID 1670 1683 BY SIMILARITY.
FT DISULFID 1689 1701 BY SIMILARITY.
FT DISULFID 1696 1710 BY SIMILARITY.
```

Query Match 12.3%; Score 724; DB 1; Length 2907;

Best Local Similarity 26.1%; Pred. No. 3,2e-40;

Matches 25; Conservative 98; Xismatches 300; Indels 314; Gaps 51;

Qy 45 DVDECAQGLDCHADALCCNTFTSYKSCCKPGYQG--EGROCEDIDEGONEUNGCVHDC 102

Db 1193 DINECSLS-DNLGRNGKCVNMGTYQCSNCPGYQATPDQGGCTDIDEC-MINWNGSCTCC 1250

Qy 103 LNTPGNVRCITPDGFMLADHGHNCLDVDECLENN-----GGQHTCVNVMNGYECCKKEG 157

Db 1251 TNSEGSVEGCSGEGYALMPDGRSCADIDEC-ENNPICIDGG---OCTNIPGEYRCLCYDG 1306

Qy 158 FFLSDNCHTCIHRSEEGISLVNKHCHCSHICK-----EAPRGSVACECRPGFELAKNORDC 213

Db 1307 FNASMDNKTCDVNECCLN-----FNICMFGECENTKGSFICHQGLGVSKKGTGC 1358



```

FT DISULFID 469 482 BY SIMILARITY.
FT DISULFID 488 499 BY SIMILARITY.
FT DISULFID 495 508 BY SIMILARITY.
FT DISULFID 510 523 BY SIMILARITY.
FT DISULFID 529 540 BY SIMILARITY.
FT DISULFID 536 549 BY SIMILARITY.
FT DISULFID 551 564 BY SIMILARITY.
FT DISULFID 570 581 BY SIMILARITY.
FT DISULFID 577 590 BY SIMILARITY.
FT DISULFID 592 605 BY SIMILARITY.
FT DISULFID 611 622 BY SIMILARITY.
FT DISULFID 618 631 BY SIMILARITY.
FT DISULFID 633 646 BY SIMILARITY.
FT CARBOHYD 221 221 N-LINKED (GICNAC...) (POTENTIAL).
FT CARBOHYD 890 890 N-LINKED (GICNAC...) (POTENTIAL).
SQ SEQUENCE 956 AA; 106779 MW; 384C22770B6E6EE CRC64;

Query Match 12.3%; Score 692; DB 1; Length 956;
Best Local Similarity 29.8%; Pred. No. 1.3e-38;
Matches 159; Conservative 72; Mismatches 224; Indels 78; Gaps 14;

QY 62 CONTPTSYKCSKPGY--QGEGRQCEDIDEQNELNGGVCHDCLNIPNYRCTCFDGMVL 119
DB 253 CLNTPGSYICKKQGYILSTDKTCRIODLCATE-DHGCEQLCVNMLGFSVCQYGYTL 311
QY 120 AHGHCNCLDVDECLNNGSCQHTCVNMGSYECCKEGFFLSDNOHTCIHRESEGLSCMN 179
DB 312 AEDGKRECTAMDYCASHGCEHCENVAESSYLRCHEGFALNSDKTKSKID----YCAS 367
QY 180 KDHCSSHCICKEAPRGVACCEPRGPELAKNQRDC--ILTCNHNNGGCGHSCDDTDADPEC 237
DB 368 SNHGCOHECVNA-QTSALCRCLGFMPLNDRKTCRRINYCALNKPGECEHCVENTEEGHVC 426
QY 238 SCHPOYKMHETDGRSCLG-----REDTVLEVTESNTTSVVG-----DKRVKRLME 284
DB 427 RCQGNLNPNGKTCRSVDHCAQDHGCEQLCLNTEESFVQCSEGF--INDLKTCSRAJ 486
QY 285 TCANVNGGCDRTCKDTSTGVHSCSPVFTLQDGGTKCDIDBCQPRNGGCDHFCKNIVGS 344
DB 487 YCLLSNHCSEYCVNTDKSFACQCEGHVLRSDGKTCALKLSGALGDHGCHESCVSSSDS 546
QY 345 FDCGCKKGLKLTDEKSCQVDEC-SJDRTCDSCHINHPGT-FACACNRYTL-YGFTHCG 402
DB 547 FVQCQFEGVILRDGKTCRRKQVQDVNHCBEHLGVNSGESVYCKSLGFLAEDGKRCR 606
QY 403 C-NECSINNGGCGQVQVNTVGSVECOCHPGYKLNKXQCV-----444
DB 607 RNVVCKSTQHGCEHMCVNVNGSYLCRSEGFVLAEDGKCRKCTEGPIDLVFVIDGSKS- 666
QY 445 -----YKGLUPT-SVSPRYSLHCGKSGGGGCFLRCHSGIHLSSDVTIRTSTV 492
DB 667 GEENFETVCHFTVGIIDS-LAVSPKA-----ARVQLQYSTQVTRTEFT 708
QY 493 FK-----LNEGKSLKVAEFPPEGLPALPEKHSVKSFRYVNLTCSSGKQVP 542
DB 709 LRGFSASAKYKXAVTHMKYMGKSGMTGLAKH-MFERSFTQEGARPPESTQV 760

RESULT 4
FBN1_BOVIN
ID FBN1_BOVIN STANDARD; PRT: 2871 AA.
AC F98133;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Fibrillin 1 precursor (NP340).
GN FBN1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID:9913;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RX MEDLINE=95137597; PubMed=7835900;
RA Tilstra D.J., Potter K.A., Byers P.H.;
RT "Sequence of the coding region of the bovine fibrillin cDNA and
RL Genomics 23:480-485(1994).";
RN [2];
RP PARTIAL SEQUENCE.
RX MEDLINE=96132851; PubMed=8557636;
RA Gibson M.A., Hatzinikolas G., Kumaratilake J.S., Sandberg L.B.,
RA Nicholl J.K., Sutherland G.R., Cleary E.G.;
RT "Further characterization of proteins associated with elastic fiber
RL J. Biol. Chem. 271:10956-10963(1996).";
CC -1- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE
CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
CC -1- PTM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER
CC FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE
CC MICROFIBRILS.
CC -1- SIMILARITY: Contains 47 EGF-like domains.
CC -1- SIMILARITY: Contains 7 TGF-beta binding protein (TGFBP) domains.
CC This SW-SS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EXBL: L28748; AAA74122...
DR PIR: A55567; A55567.
DR HSP: P35555; IAPU.
DR InterPro: IPR00152; Asx hydroxyl.
DR InterPro: IPR001881; EGF_CA.
DR InterPro: IPR006269; EGF_Like.
DR InterPro: IPR002212; Fibril-associ.
DR Pfam: PF00008; EGF; 46.
DR Pfam: PF00683; TB; 9.
DR SMART: SM00179; EGF_CA; 42.
DR PROSITE: PS00010; ASX HYDROXYL; 43.
DR PROSITE: PS00022; EGF_Like; 2.
DR PROSITE: PS01186; EGF_Like; 38.
DR PROSITE: PS01187; EGF_CA; 43.
DR Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KW Repeat; Signal; Multigene family.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 2871 FIBRILLIN 1.
FT DOMAIN 81 112 EGF-LIKE 1.
FT DOMAIN 115 146 EGF-LIKE 2.
FT DOMAIN 147 178 EGF-LIKE 3.
FT DOMAIN 246 287 EGF-LIKE 4.
FT DOMAIN 288 329 EGF-LIKE 5.
FT REPEAT 330 390 TGFBP 1.
FT DOMAIN 392 446 PRO-RICH.
FT DOMAIN 449 489 EGF-LIKE 6.
FT DOMAIN 490 529 EGF-LIKE 7.
FT DOMAIN 530 571 EGF-LIKE 8.
FT DOMAIN 572 612 EGF-LIKE 9.
FT DOMAIN 613 653 EGF-LIKE 10.
FT REPEAT 654 722 TGFBP 2.
FT DOMAIN 723 764 EGF-LIKE 11.
FT REPEAT 765 806 EGF-LIKE 12.
FT DOMAIN 807 846 EGF-LIKE 13.
FT DOMAIN 847 952 EGF-LIKE 14.
FT REPEAT 952 1027 TGFBP 3.
FT DOMAIN 1028 1069 EGF-LIKE 15.
FT DOMAIN 1070 1112 EGF-LIKE 16.
FT DOMAIN 1113 1154 EGF-LIKE 17.
FT DOMAIN 1155 1196 EGF-LIKE 18.
FT DOMAIN 1197 1237 EGF-LIKE 19.

```

FT	DISULPID	937	950	BY SIMILARITY.
FT	DISULPID	1032	1044	BY SIMILARITY.
FT	DISULPID	1039	1053	BY SIMILARITY.
FT	DISULPID	1055	1068	BY SIMILARITY.
FT	DISULPID	1074	1086	BY SIMILARITY.
FT	DISULPID	1081	1095	BY SIMILARITY.
FT	DISULPID	1097	1111	BY SIMILARITY.
FT	DISULPID	1117	1129	BY SIMILARITY.
FT	DISULPID	1124	1138	BY SIMILARITY.
FT	DISULPID	1140	1153	BY SIMILARITY.
FT	DISULPID	1159	1171	BY SIMILARITY.
FT	DISULPID	1166	1180	BY SIMILARITY.
FT	DISULPID	1182	1195	BY SIMILARITY.
FT	DISULPID	1201	1212	BY SIMILARITY.
FT	DISULPID	1208	1221	BY SIMILARITY.
FT	DISULPID	1223	1236	BY SIMILARITY.
FT	DISULPID	1242	1254	BY SIMILARITY.
FT	DISULPID	1249	1263	BY SIMILARITY.
FT	DISULPID	1265	1278	BY SIMILARITY.
FT	DISULPID	1284	1296	BY SIMILARITY.
FT	DISULPID	1291	1305	BY SIMILARITY.
FT	DISULPID	1307	1321	BY SIMILARITY.
FT	DISULPID	1326	1339	BY SIMILARITY.
FT	DISULPID	1333	1348	BY SIMILARITY.
FT	DISULPID	1350	1361	BY SIMILARITY.
FT	DISULPID	1367	1380	BY SIMILARITY.
FT	DISULPID	1374	1389	BY SIMILARITY.
FT	DISULPID	1391	1402	BY SIMILARITY.
FT	DISULPID	1409	1420	BY SIMILARITY.
FT	DISULPID	1415	1429	BY SIMILARITY.
FT	DISULPID	1431	1444	BY SIMILARITY.
FT	DISULPID	1450	1461	BY SIMILARITY.
FT	DISULPID	1456	1470	BY SIMILARITY.
FT	DISULPID	1472	1485	BY SIMILARITY.
FT	DISULPID	1491	1502	BY SIMILARITY.
FT	DISULPID	1497	1511	BY SIMILARITY.
FT	DISULPID	1513	1526	BY SIMILARITY.
FT	DISULPID	1610	1622	BY SIMILARITY.
FT	DISULPID	1617	1631	BY SIMILARITY.
FT	DISULPID	1633	1646	BY SIMILARITY.
FT	DISULPID	1652	1663	BY SIMILARITY.

Query Xatch 12.3%; Score 691; DB 1; Length: 2871;

Best Local Similarity 22.7%; Pred. No. 5e-38;

Matches 271; Conservative 107; Mismatches 375; Indels 440; Gaps 49;

QY	45	DYDECAQGLDDCHADALCONTFTSYKSCCKPQVQEGROCEFD	DECGNENLGGCVHDCLN	154
QY	1363	DIDESGNGTHKCSQHADCKNTWGSFRLCKEGYTGDFCTCTD	DEKSENLCGNQCCLN	1422
QY	105	IPGNVCTCFDGFMLAHGHNCLDVDC	-----LENNG	137
QY	1423	APGYRCECDMGFVPSADGKACED-DECSLPNICVPGTCHNLPGLFRCEIGYELDRSG	1482	
QY	138	G-----CQHCVNVMGSEYCCCKEGEFLSDNCHTCIH-----	ES	171
QY	1493	GNCTDVNECLDPTTCISGNCVNTPGSYTCDCCPPDFELNTRVGCVTGRSGNCYLDIRPRG	1542	
QY	172	EEG-LSCMK-----KRGCSHI-----	187	
QY	1543	DNGDTACSNEIGVGVSKASCCSLGKAWCTPCELCFPVNTSEYKILPCGEGFRPNPITV	1602	
QY	188	-----CKEAPR-----GSVACECRPGELAKNQDCC--ILTCNH-----	GNCG	223
QY	1603	1-EDIDECEGLGLCOGSKCINTFGSFQCRCTGYVNLNEDTRVCODWNECETPGICGPG-	1661	
QY	224	CQHSQCDTADGFECSCHPYKMHTRDGRSCLEREDTVLEVTESNTTSWQDGKRVKRLLM	283	
QY	1662	-----TCVNTVGNVTCICPPDYMQVQGNQCNQWRSICVYNYIADNQTCCGELLFVTKKM	1718	
QY	284	ETCAVNG-GCDRTK-----DTSGV-----	304	
QY	1719	CCCSNIGRANKRPFQCEIPSTDEFEATCGSORPGFVIDIYTGLPVDR-DECRETFGVC	1778	

```
QY 305 -----HCSPVGTFLDQKTKDIDBCQ----- 328
Db 1779 NGVCINMWSFRCECPVGFYNDKLLUCEDIDECQNGPVCQNAECINTAGSYRCCCKPG 1838
QY 329 ---TRNGGCDH-----FCNKIVGDFCGCKKFKLLTDEKSCQDVDECSLD 371
Db 1839 YRFTSTGQCDNRNECOEIPNCSHGOCIDTVGSFYCLHTGKTNADQTMCLDINECED 1898
QY 372 RTCDHSCINHPGTFACNRYTLXGTHCGDTNCSINNGG-CQO-VCVNIVGSYECQC 429
Db 1899 ACNGTCRNTIGSFNCRNHFILSHNDICIDVDECATGNLCRNGQCINTVGSFQCQC 1958
QY 430 HPGYKLNHKKOCVEYKGLL---PTSVPRVSLHCKSGGGDGLRCHSGIHLSDVTT 487
Db 1959 NEGVEVAPDGRICVDINECLLPKRCAPGT---CQNLDSYRCI---CPFGVSLQNDKCE- 2012
QY 488 RTSVTFKLN-----GKCSLKNAE-----LPPEGLR-PALPEKHSSVKESFRYN 531
Db 2013 -----DIDECVEEPEICALGTCS---NTGSPFKCLCPDGFSLSTGRCCQLRKYCYAK 2064
QY 532 LTCSSKQVPGARPRSTKEMFIVPELETNQKEVTASCLSL-----CIVKRTSKR 584
Db 2065 F-----EGKCSSPK-----SRHSGKQCCCKALKGEGMDPCCLCPTF- 2103
QY 585 LRKAIRTLKAVHREOFHLQL-----SGMNLVAKKPPRTSERQAES----- 626
Db 2104 -----DEAFRIQCPYQSGIIVGDDSAVDECKEFDVKHGGCINTDOSY 2149
QY 627 ---CGVQGHAEQVQ---SCRAGTYOGARERCILCPNGTQNEBQWTCPEPPRPGNS 68C
Db 2150 RCECPGYILOGNECVDTDECSVGN-----PCNGTKCKNVIGGPEC-TCBEGSEP 2198
QY 681 GALKTEPANNMECCG---LCQPGESADGAPCOLCALGTQPEAGRTSC---FPCGGG 734
Db 2199 GPMTCCE--DINECAONPLLCAFRCNTVGSYECK-CPAGYVLRE-DRMCKOEDECEG 2254
QY 735 LAKHOGATSFODCEFRVQ-----CSPGHFYNTTHRCIRCVPVGTQPEFGKNCVSCPG 789
Db 2255 ---KHCEAKQWECKNL-GTYL-CICPGVQRRPDGEGCV-----DENECQCKPG 2320
QY 790 NTTDFEGSTNTQCKNRCCGELGDT-----GYIESPNYGNYPANTECTWTIN?PPK 844
Db 2301 -----ICENGRCLNTRGTYCECDGETASPN----- 2327
QY 845 RRLIVVPEILFIEDDCG-----YLVVRKTSNNSVTVETCTQYER--- 886
Db 2328 -----QDECLDNREGYCFTEVLQNMCOIGSSNRPVTKSECCCGRGKWP 2373
QY 889 -----PIAFTRSKLWLOFKSNEGNSARGFQVPVYVYEDYQELLEDIVRDG 936
Db 2374 HCEICPFQGTVAFKLCL-----PHGRGFTNGAD:DE--CKVIHDVCRNG 2416
```

## RESULT 5

```
FBN1_HUMAN
ID FBN1_HUMAN STANDARD; PRT; 2871 AA.
AC P35555;
DT 01-JUN-1994 (Rel. 29, Created)
DI 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Fibrillin 1 precursor.
GN FBN1 OR FBN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=93372860; PubMed=8364578;
RA Pereira L.V., D'Alessio M., Ramirez F., Lynch J.R., Sykes B.,
RA Pangilinan T., Bonadio J.;
```

```
RT "Genomic organization of the sequence coding for fibrillin, the
RL defective gene product in Marfan syndrome.";
RN Hum. Mol. Genet. 2:961-968(1993).
RP [2]
RP SEQUENCE OF 1-932 FROM N.A.
RC T-SSUE=Fibroblast, and Placenta;
RX MEDLINE=9401C947; PubMed=7691719;
RA Carson G.M., Chabberg S.C., Dietz H.C., Charbonneau N.J., Sakai L.Y.;
RT "Fibrillin binds calcium and is coded by cDNAs that reveal a
RL multidomain structure and alternatively spliced exons at the 5'
end.";
RN Genomics 17:476-484(1993).
RP [3]
RP SEQUENCE OF 899-2871 FROM N.A.
RX MEDLINE=91304567; PubMed=8522206;
RA Maslen C.L., Carson G.M., Maddox B.K., Glanville R.W., Sakai L.Y.;
RC "Partial sequence of a candidate gene for the Marfan syndrome.";
RL Nature 352:334-337(1991).
RP [4]
RP SEQUENCE OF 813-1313 FROM N.A.
RX MEDLINE=91304567; PubMed=8522206;
RA Lee S., Godfrey M., Vitale E., Hori H., Mattei M.-G., Sarfarazi M.,
RA Tsipouras P., Ramirez F., Hollister C.W.;
RT "Linkage of Marfan syndrome and a phenotypically related disorder to
RL two different fibrillin genes.";
RN Nature 352:330-334(1991).
RP [5]
RP CHARACTERIZATION.
RX MEDLINE=93317849; PubMed=860873;
RA Sakai L.Y., Keene D.R., Glanville R.W., Bachinger H.P.;
RT "Purification and partial characterization of fibrillin, a cysteine-
RL rich structural component of connective tissue microfibrils.";
RN J. Biol. Chem. 266:14763-14770(1991).
RP [6]
RP STRUCTURE BY NMR OF 2054-2125.
RX MEDLINE=96031893; PubMed=8562480;
RA Yuan X., Downing A.K., Knott V., Handford P.A.;
RT "Solution structure of the transforming growth factor beta-binding
RL protein-like module, a domain associated with matrix fibrils.";
RN EMBO J. 16:6659-6666(1997).
RP [7]
RP STRUCTURE BY NMR OF 2124-2205.
RX MEDLINE=96174615; PubMed=8594563;
RA Knott V., Downing A.K., Cady C.M., Handford P.A.;
RT "Calcium binding properties of an epidermal growth factor-like domain
RL pair from human fibrillin-1.";
RN J. Mol. Biol. 255:22-27(1996).
RP [8]
RP STRUCTURE BY NMR OF 2124-2205.
RX MEDLINE=96222301; PubMed=8653794;
RA Downing A.K., Knott V., Werner J.M., Cady C.M., Campbell I.D.,
RA Handford P.A.;
RT "Solution structure of a pair of calcium-binding epidermal growth
RL factor-like domains: implications for the Marfan syndrome and other
RL genetic disorders.";
RN Cell 85:597-605(1996).
RP [9]
RP REVIEW ON MFS VARIANTS.
RX MEDLINE=96174615; PubMed=8594563;
RA Colod G., Beroud C., Soussi T., Junien C., Boileau C.;
RT "Software and database for the analysis of mutations in the human
RL FBN1 gene.";
RN Nucleic Acids Res. 24:137-141(1996).
RP [10]
RP REVIEW ON MFS VARIANTS.
RX MEDLINE=97169383; PubMed=9016526;
RA Colod-Beroud G., Beroud C., Ades L., Slack C., Boxer M., Brock D.C.,
RA Godfrey X., Hayward C., Karttunen L., Milewicz D., Peltonen L.,
RA Richards R.I., Wang W., Junien C., Boileau C.;
RT "Marfan Database (second edition): software and database for the
RL analysis of mutations in the human FBN1 gene.";
RN Nucleic Acids Res. 25:147-150(1997).
RN [11]
```





QY 329 -----TRNGGCOH-----FCNIVGSPDCCKGKFKLLTDEKSCQVDECSL 370  
 Db 1838 GYRFTSTQCNDNRNCOBIPNICSHGQCIDTVGSPYCLCHTGFKNDDQTMCLDINCEK 1837  
 QY 371 DRTCHSCINHPGTACACNRGYTLGTHCGDTNECSINNG-CCQ-VCVNTVGSVECCQ 428  
 Db 1898 DACNGTORTNTGSCNCRNGHFLSHNNDICIDVDECASGNLCRNGCCINTVGSFQCC 1957  
 QY 429 CHPGYKHLNKKDCVVEKGLL--PTSVPRVSLHCGKSGGGGDFLCHSGIHLSSDVT 486  
 Db 1958 CNEGYEAPDGRCTVDINECLLEPKCAPGT--CQNLDSGYRCI--CPCGYSLONEKCE 2012  
 QY 487 IRTSVTFKLINE-----GKSLKNAE-----LFPEG-R-PALPEKSHVSEFRV 530  
 Db 2013 -----DIDECVERPEICALGTCS--NTEGSKLCPEGFSLSGSSRRCCDLMSVCYA 2063  
 QY 531 NLTCSSGQKVPAGRPSTPKEMFTVFELETNCKEVTASCDLS-----CIVKRTK 583  
 Db 2064 KF-----EGKCSPK-----SRHSHKQECCECCALKGEGMGDFCBLCTEP 2103  
 QY 584 RLRKAIRTLRKAVHREQHLQ-----SGYNLDVAKKPPRTSEROAES----- 626  
 Db 2104 -----DEAFRQICPYGSGIIVGPDSDAVDMDECKEPDVCCKHGQCINTDGS 2148  
 QY 627 -----CGVGQGHAEKCV-----LCQGEYSADGAPQCLCALGTFCPEAGRTSC--FPCGG 733  
 Db 2198 PGPMMTCE--DINECAQNPLLCARFVNTYGSYECK-CPVGVVLE-DRMCKDEDECEE 2253  
 QY 734 G-ATKHQGA:SFQDCETRVQ-----CSPGHFVNTTHRCIRCPVGTGYOPEFGKNCVSCP 788  
 Db 2254 G---KHDCTEKOMECKNLJGTMYCICPGYQRRPDGECV-----DENECQTKP 2299  
 QY 789 GNTTTFDGSNTITOCNRRCGELGDET-----GVIESPNYPGNYPANTCTWTINPP 843  
 Db 2300 G-----ICENGRCLNTRGSYTCECNDGFTASPK----- 2327  
 QY 844 KRRILIVVPEIFLEDCCGD-----YLVMKRTSSNSVTVYTCQTVR--- 888  
 Db 2328 -----QDECLDNREGYCFTEVLQNMCIQSSNRNPTVTKSECCDGRGCG 2372  
 QY 889 -----PIAFTSRKKLWTFKNSARGQVFPVYDEYQQLIDIDVUDG 936  
 Db 2373 PHCEICPFOGTVAFKKLC-----PHORGFMINGADIDE--CKVIHDVCNKG 2416  
 RESULT 6  
 FBN1\_MOUSE STANDARD; PRT: 2871 AA.  
 AC Q61554; Q60826;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Fibrillin 1 precursor.  
 GN FBN1 OR FBN-1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95130561; PubMed=78299516;  
 RA Yin W., Gernier J., Sanguinetti C., Smiley S., Pangilinan T.,  
 RA Pereira L., Ramirez F., Bonadio J.,  
 RT "Primary structure and developmental expression of Fbn-1, the mouse  
 RT fibrillin gene."  
 RL J. Biol. Chem. 270:1798-1806(1995).  
 RN (2)  
 RP SEQUENCE FROM N.A.

RC STRAIN=CD-1; TISSUE=Kidney;  
 RA Cta K., Kumar A., Wada J., Liu Z., Kanwar Y.S.;  
 RL Submitted (APR-1995) to the EMBL/Genbank/DBSD databases.  
 CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS  
 CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE  
 CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.  
 CC -!- PTM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER  
 CC FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE  
 CC MICROFIBRILUS (BY SIMILARITY).  
 CC -!- SIMILARITY: Contains 47 EGF-like domains.  
 CC -!- SIMILARITY: Contains 7 TGF-beta binding protein (TGF-beta) domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR ENBL; L29454; AAA56840.1; --  
 DR ENBL; U22493; AAA64227.1; --  
 DR PIR; A55624; A55624.  
 DR HSP; P35555; IAPU.  
 DR MGC; MG1:95489; Fbn1.  
 DR InterPro; IPR000152; Asx\_HydroxyL.  
 DR InterPro; IPR001188; EGF\_Ca  
 DR InterPro; IPR006209; EGF\_Like.  
 DR InterPro; IPR002212; Fibril-assoc.  
 DR Pfam; PF00038; EGF; 46.  
 DR Pfam; PF0683; TB; 9.  
 DR SMART; SM00179; EGF\_CA; 42.  
 DR PROSITE; PS00310; ASX\_HYDROXYL; 43.  
 DR PROSITE; PS00322; EGF\_1; 2.  
 DR PROSITE; PS01186; EGF\_2; 38.  
 DR PROSITE; PS01187; EGF\_CA; 43.  
 KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;  
 KW Repeat; Signal; Multigene family.  
 FT SIGNAL 1 27 POTENTIAL  
 FT CHAIN 28 2871 FIBRILLIN 1.  
 FT DOMAIN 81 112 EGF-LIKE 1.  
 FT DOMAIN 115 146 EGF-LIKE 2.  
 FT DOMAIN 147 178 EGF-LIKE 3.  
 FT DOMAIN 246 287 EGF-LIKE 4.  
 FT DOMAIN 288 329 EGF-LIKE 5.  
 FT DOMAIN 330 401 TGF-beta 1.  
 FT DOMAIN 402 446 PRO-RICH.  
 FT DOMAIN 449 489 EGF-LIKE 6.  
 FT DOMAIN 493 529 EGF-LIKE 7.  
 FT DOMAIN 530 571 EGF-LIKE 8.  
 FT DOMAIN 572 612 EGF-LIKE 9.  
 FT DOMAIN 613 653 EGF-LIKE 10.  
 FT DOMAIN 656 721 TGF-beta 2.  
 FT DOMAIN 723 764 EGF-LIKE 11.  
 FT DOMAIN 765 806 EGF-LIKE 12.  
 FT DOMAIN 807 846 EGF-LIKE 13.  
 FT DOMAIN 910 951 EGF-LIKE 14.  
 FT DOMAIN 952 1018 TGF-beta 3.  
 FT DOMAIN 1020 1069 EGF-LIKE 15.  
 FT DOMAIN 1070 1112 EGF-LIKE 16.  
 FT DOMAIN 1113 1154 EGF-LIKE 17.  
 FT DOMAIN 1155 1196 EGF-LIKE 18.  
 FT DOMAIN 1197 1237 EGF-LIKE 19.  
 FT DOMAIN 1238 1279 EGF-LIKE 20.  
 FT DOMAIN 1280 1321 EGF-LIKE 21.  
 FT DOMAIN 1322 1362 EGF-LIKE 22.  
 FT DOMAIN 1363 1403 EGF-LIKE 23.  
 FT DOMAIN 1404 1445 EGF-LIKE 24.  
 FT DOMAIN 1446 1486 EGF-LIKE 25.  
 FT DOMAIN 1487 1527 EGF-LIKE 26.  
 FT DOMAIN 1528 1569 TGF-beta 4.  
 FT DOMAIN 1606 1647 EGF-LIKE 27.  
 FT DOMAIN 1648 1688 EGF-LIKE 28.

```
FT DOMAIN 1689 1758 TGFBR 5. 1758 1758 BY SIMILARITY.
FT DOMAIN 1766 1807 EGF-LIKE 29, CALCIUM-BINDING.
FT DOMAIN 1808 1848 EGF-LIKE 30, CALCIUM-BINDING.
FT DOMAIN 1849 1889 EGF-LIKE 31, CALCIUM-BINDING.
FT DOMAIN 1891 1929 EGF-LIKE 32, CALCIUM-BINDING.
FT DOMAIN 1930 1972 EGF-LIKE 33, CALCIUM-BINDING.
FT DOMAIN 1973 2012 EGF-LIKE 34, CALCIUM-BINDING.
FT DOMAIN 2013 2054 EGF-LIKE 35, CALCIUM-BINDING.
FT DOMAIN 2055 2121 TGFBR 6.
FT DOMAIN 2122 2165 EGF-LIKE 36, CALCIUM-BINDING.
FT DOMAIN 2166 2205 EGF-LIKE 37, CALCIUM-BINDING.
FT DOMAIN 2206 2246 EGF-LIKE 38, CALCIUM-BINDING.
FT DOMAIN 2247 2290 EGF-LIKE 39, CALCIUM-BINDING.
FT DOMAIN 2291 2332 EGF-LIKE 40, CALCIUM-BINDING.
FT DOMAIN 2333 2400 TGFBR 7.
FT DOMAIN 2401 2443 EGF-LIKE 41, CALCIUM-BINDING.
FT DOMAIN 2444 2484 EGF-LIKE 42, CALCIUM-BINDING.
FT DOMAIN 2485 2523 EGF-LIKE 43, CALCIUM-BINDING.
FT DOMAIN 2524 2566 EGF-LIKE 44, CALCIUM-BINDING.
FT DOMAIN 2567 2606 EGF-LIKE 45, CALCIUM-BINDING.
FT DOMAIN 2607 2647 EGF-LIKE 46, CALCIUM-BINDING.
FT DOMAIN 2648 2687 EGF-LIKE 47, CALCIUM-BINDING.
FT DISULFID 85 94 BY SIMILARITY.
FT DISULFID 89 100 BY SIMILARITY.
FT DISULFID 102 111 BY SIMILARITY.
FT DISULFID 119 129 BY SIMILARITY.
FT DISULFID 123 134 BY SIMILARITY.
FT DISULFID 136 145 BY SIMILARITY.
FT DISULFID 150 160 BY SIMILARITY.
FT DISULFID 154 166 BY SIMILARITY.
FT DISULFID 168 177 BY SIMILARITY.
FT DISULFID 250 262 BY SIMILARITY.
FT DISULFID 257 271 BY SIMILARITY.
FT DISULFID 273 286 BY SIMILARITY.
FT DISULFID 292 304 BY SIMILARITY.
FT DISULFID 299 313 BY SIMILARITY.
FT DISULFID 315 328 BY SIMILARITY.
FT DISULFID 453 465 BY SIMILARITY.
FT DISULFID 460 474 BY SIMILARITY.
FT DISULFID 476 488 BY SIMILARITY.
FT DISULFID 494 504 BY SIMILARITY.
FT DISULFID 499 513 BY SIMILARITY.
FT DISULFID 515 528 BY SIMILARITY.
FT DISULFID 534 546 BY SIMILARITY.
FT DISULFID 541 555 BY SIMILARITY.
FT DISULFID 557 570 BY SIMILARITY.
FT DISULFID 576 587 BY SIMILARITY.
FT DISULFID 582 596 BY SIMILARITY.
FT DISULFID 598 611 BY SIMILARITY.
FT DISULFID 617 628 BY SIMILARITY.
FT DISULFID 623 637 BY SIMILARITY.
FT DISULFID 639 652 BY SIMILARITY.
FT DISULFID 727 739 BY SIMILARITY.
FT DISULFID 734 748 BY SIMILARITY.
FT DISULFID 750 763 BY SIMILARITY.
FT DISULFID 769 781 BY SIMILARITY.
FT DISULFID 776 790 BY SIMILARITY.
FT DISULFID 792 805 BY SIMILARITY.
FT DISULFID 811 821 BY SIMILARITY.
FT DISULFID 816 830 BY SIMILARITY.
FT DISULFID 832 845 BY SIMILARITY.
FT DISULFID 914 926 BY SIMILARITY.
FT DISULFID 921 935 BY SIMILARITY.
FT DISULFID 937 950 BY SIMILARITY.
FT DISULFID 1032 1044 BY SIMILARITY.
FT DISULFID 1039 1053 BY SIMILARITY.
FT DISULFID 1055 1068 BY SIMILARITY.
FT DISULFID 1074 1086 BY SIMILARITY.
FT DISULFID 1081 1095 BY SIMILARITY.
FT DISULFID 1097 1111 BY SIMILARITY.
FT DISULFID 1117 1129 BY SIMILARITY.
FT DISULFID 1124 1138 BY SIMILARITY.
FT DISULFID 1140 1153 BY SIMILARITY.
FT DISULFID 1159 1171 BY SIMILARITY.
FT DISULFID 1166 1180 BY SIMILARITY.
FT DISULFID 1182 1195 BY SIMILARITY.
FT DISULFID 1201 1212 BY SIMILARITY.
FT DISULFID 1223 1236 BY SIMILARITY.
FT DISULFID 1238 1254 BY SIMILARITY.
FT DISULFID 1249 1263 BY SIMILARITY.
FT DISULFID 1265 1278 BY SIMILARITY.
FT DISULFID 1284 1296 BY SIMILARITY.
FT DISULFID 1291 1305 BY SIMILARITY.
FT DISULFID 1327 1339 BY SIMILARITY.
FT DISULFID 1326 1348 BY SIMILARITY.
FT DISULFID 1350 1361 BY SIMILARITY.
FT DISULFID 1367 1380 BY SIMILARITY.
FT DISULFID 1374 1389 BY SIMILARITY.
FT DISULFID 1391 1402 BY SIMILARITY.
FT DISULFID 1408 1420 BY SIMILARITY.
FT DISULFID 1415 1429 BY SIMILARITY.
FT DISULFID 1431 1444 BY SIMILARITY.
FT DISULFID 1450 1461 BY SIMILARITY.
FT DISULFID 1456 1470 BY SIMILARITY.
FT DISULFID 1472 1485 BY SIMILARITY.
FT DISULFID 1491 1502 BY SIMILARITY.
FT DISULFID 1497 1511 BY SIMILARITY.
FT DISULFID 1513 1526 BY SIMILARITY.
FT DISULFID 1610 1622 BY SIMILARITY.
FT DISULFID 1617 1631 BY SIMILARITY.
FT DISULFID 1633 1646 BY SIMILARITY.
FT DISULFID 1652 1663 BY SIMILARITY.
FT DISULFID 1658 1672 BY SIMILARITY.
FT DISULFID 1674 1687 BY SIMILARITY.
Query Match 12.1% Score 690; DB 1; Length 2871;
Best Local Similarity 23.4% Pred No. 5.8e-38;
Matches 276; Conservative 104; Mismatches 384; Indels 418; Gaps 51;
QY 45 DVDECAQGLDDCHADALCONTPSYKSCXKPGYQGEGRQCEIDECGNELNGCGVHCLN 104
DB 1363 DLDECSNGTHMCSQHADCKNTMGSYRCLCKDGVGTGDTCTDLDECSNLCNGGCLN 1422
QY 105 IPGNRYCTCDGFMALHDGHNCILDVDC -----LENG 137
DC 1423 APGGYRCEDYGFVPSADGKACEDIDECSPNTCVFOTCHNLPGLFRCEIEGYELDRSG 1482
QY 138 G-----CGHTCVNMGSYECCKEGLFSENCHTCIH-----RS 171
DB 1483 GNCTDVNECLDPTTC:SGNCVNTPGSYTDCSDPDELNPTRVGCVDRSNCVNIAPRG 1542
QY 172 EGLSCVNRKHG-----CSHI----- 187
DB 1543 ENGDTACSNEIGVGVSASCCSLGKAWGTFCBLCPSPVNTSEYKILCPGGEGRFPNPTV 1602
QY 188 -----CKEAPR-----GSVACEGRGPELAKNQRDC--ILTCNH-----QNGG 223
DB 1603 ILEDIDECQELPGLCGGKGINTFGQCRCPTGYJNEJTRVCDVNEGETPGICGPG- 1661
QY 224 CQHSCTDTADGPECSCHPOYKMHITDGRASCLEREDTLEVTESNTTSVVDGDKRVRKRLM 283
DB 1662 ---TCYNTVGNVTC:CPDYMQVNGNNGNNDKRSICRYNYADNCTCQGLLFWMTKKY 1718
QY 284 ETCAN-----NGGCRCTCKDTT----- 302
DB 1719 CCSSYNIRRAWNKPEEQ-CPIPSDEPATLCSORPGFVIDYTGLPVDIDECREIPGVC 1777
QY 303 --GV-----HCSQVGGTLOLDGKTCXGIDECQ----- 328
DB 1778 ENGVC:NMVGSFRCFCVGFYNDKLLVCEDIDECQNGPYCLRNAECINPAGSYRCCKP 1837
QY 329 ----TRNGGCDH-----FCKNIVGSFDCGCKKFKJLTDEKSCQVDECSJ 370
DB 1838 GYRJ:STGQCNDRNECQEIPIKSHGOCIDTVSGFYCLCHTGFTNFQDQNCUSINECEP 1897
```



```

FT DISULFID 406 417 BY SIMILARITY.
FT DISULFID 413 426 BY SIMILARITY.
FT DISULFID 428 441 BY SIMILARITY.
FT DISULFID 447 458 BY SIMILARITY.
FT DISULFID 454 467 BY SIMILARITY.
FT DISULFID 469 482 BY SIMILARITY.
FT DISULFID 488 499 BY SIMILARITY.
FT DISULFID 495 508 BY SIMILARITY.
FT DISULFID 513 523 BY SIMILARITY.
FT DISULFID 529 540 BY SIMILARITY.
FT DISULFID 536 549 BY SIMILARITY.
FT DISULFID 551 564 BY SIMILARITY.
FT DISULFID 561 581 BY SIMILARITY.
FT DISULFID 577 590 BY SIMILARITY.
FT DISULFID 592 605 BY SIMILARITY.
FT DISULFID 611 622 BY SIMILARITY.
FT DISULFID 618 631 BY SIMILARITY.
FT DISULFID 633 646 BY SIMILARITY.
FT CARBOHYD 221 231 N-LINKED (GLUCAC. . .) (POTENTIAL).
FT CARBOHYD 890 890 N-LINKED (GLUCAC. . .) (POTENTIAL).
FT VARSPLIC 861 879 Missing (in isoform Short).
FT CONFLICT 356 356 /FTID=VSP_001399.
FT CONFLICT 594 594 K -> E (IN REF. 3).
FT CONFLICT 644 644 V -> E (IN REF. 3).
FT CONFLICT 755 755 R -> G (IN REF. 2).
FT CONFLICT 935 935 F -> L (IN REF. 3).
FT CONFLICT 935 935 L -> F (IN REF. 2).
SQ SEQUENCE 956 AA; 106840 MW; 826B7E347178FC80 CRC64;

Query Match 12.2%; Score 684.5; DB 1; Length 956;
Best Local Similarity 32.3%; Pred. No. 4e-38;
Matches 146; Conservative 61; Mismatches 186; Indels 59; Gaps 13;

QY 62 QNTPTSYKCSKPGY--QGEGRQCEDEIDECNELNGCVHDCNLNIPGNYRCTCFDGEWL 119
Db 253 CINIPGSYVCRCKQYILNSDTCRIQDLCAE-DHNCCEQLCVNVPGSFVCCYSGVAL 311
QY 120 AUDGHNCLDVBCELENNGGCQHTCVNMGYSCKCKEGFFLSDNQHCTIHRSEGLSCWN 179
Db 312 AEDGKRCVAVDYCASENHGCHECVNADSGYLQCCHGFPALNPDKTKTKID----YCAS 367
QY 180 KDHGSHICEAPRGSVACEPRGPELAKNORDC--ILTCNHNGGCGHSCDDTDAGPEC 237
Db 368 SHHQOHECVNTD-DYSCHCLKLGFTLPNDPKTKRINVCALNKPGECEHCNVMEESYYC 426
QY 238 SCHPOYKMHITDGRSLEREDTVLEVTESNTTSVVDGKFKVRRLMETCAVNGGCDRTC 297
Db 427 RCHRGYTLDPNGKTCR-----VDHCAQODHGCQGLC 458
QY 298 KOTSTGVHSCVPVGFLLQDLGKTCXIDECOTRNGGCDHFCXN-VGSPDGGCKKGFLLT 357
Db 459 LNTEDSFVQCSEGFLLNEDLTKCSRVDYCLLSDHGCEYSCVMMDRSFACQCPGHVLR 518
QY 358 DEKSCQDVBCESL-DRTCDHSCINHPGTFACACNRYTLY--GFTHCGDTNCSINNGGC 414
Db 519 DGKTCALSKALGDHGCSEHSCVSSDSFVCCQFEGYLLREDKT-CRRKVCQAIIDHGC 577
QY 415 QVCVNTVGSYECQCHPGYKLMNKKDCVEVGLLPTSVSRVSLHCGKS-----GGGQCC 470
Db 578 EHICVNSDDSYTCECLVGFRLAEDGKRC-RRKDVCK-----STHGCHEICVNGKSY 629
QY 471 FLRCHSGIHLSSDVTIRTSTFKLNEGKCSL 502
Db 630 ICKSEGFVLAEDGRCK-----KCTEGPIDL 656

RESULT 8
FBN1_PIG STANDARD; PRT; 2871 AA.
AC Q9TV36;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)

```

```

DE Fibrillin 1 precursor.
GN FBN1.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9923;
RN [1]
RP SEQUENCE FROM N.A.
RC T-SSU=Lung;
RX MEDLINE=99156858; PubMed=10036187;
RA Biery N.A.; Eldadah Z.A.; Moore C.S.; Stetten G.; Spencer F.;
RA Dietz H.C.;
RT "Revised genomic organization of FBN1 and significance for regulated
  gene expression.";
RJ Genomics 56:70-77(1999).
CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
  THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE
  LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
CC -!- PFM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER
  MICROFIBRILS.
CC -!- SIMILARITY: Contains 47 EGF-like domains.
CC -!- SIMILARITY: Contains 7 TGF-beta binding protein (TGFBP) domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EXBL; AF073800; AAD50328.1; .
CC HSP; P35555; IAPC.
CC InterPro: IPR000152; Asx hydroxyl.
CC InterPro: IPR001881; EGF-Ca.
CC InterPro: IPR006209; EGF-like.
CC InterPro: IPR006212; Fibril-associ.
CC Pfam; PF00008; EGF; 45.
CC Pfam; PF00683; TB; 9.
CC SMART; SM00179; EGF CA; 40.
CC PROSITE; PS00010; ASX HYDROXYL; 41.
CC PROSITE; PS00022; EGF_1; 2.
CC PROSITE; PS01186; EGF_2; 36.
CC PROSITE; PS01187; EGF CA; 41.
CC Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
  Repeat; Signal; Multigene family.
CC SIGNAL; 27 POTENTIAL.
CC CHAIN; 28 2871 FIBRILLIN 1.
CC DOMAIN; 81 112 EGF-LIKE 1.
CC DOMAIN; 115 146 EGF-LIKE 2.
CC DOMAIN; 147 178 EGF-LIKE 3.
CC DOMAIN; 246 287 EGF-LIKE 4. CALCIUM-BINDING.
CC DOMAIN; 288 329 EGF-LIKE 5. CALCIUM-BINDING.
CC REPEAT; 330 390 TGFBP 1.
CC DOMAIN; 392 446 PRO-RICH.
CC DOMAIN; 449 489 EGF-LIKE 6.
CC DOMAIN; 490 529 EGF-LIKE 7. CALCIUM-BINDING.
CC DOMAIN; 530 571 EGF-LIKE 8. CALCIUM-BINDING.
CC DOMAIN; 572 613 EGF-LIKE 9. CALCIUM-BINDING.
CC DOMAIN; 613 653 EGF-LIKE 10. CALCIUM-BINDING.
CC REPEAT; 654 722 TGFBP 2.
CC DOMAIN; 723 764 EGF-LIKE 11. CALCIUM-BINDING.
CC DOMAIN; 765 806 EGF-LIKE 12. CALCIUM-BINDING.
CC DOMAIN; 807 846 EGF-LIKE 13. CALCIUM-BINDING.
CC DOMAIN; 910 951 EGF-LIKE 14. CALCIUM-BINDING.
CC REPEAT; 952 1027 TGFBP 3.
CC DOMAIN; 1028 1069 EGF-LIKE 15. CALCIUM-BINDING.
CC DOMAIN; 1070 1112 EGF-LIKE 16. CALCIUM-BINDING.
CC DOMAIN; 1113 1154 EGF-LIKE 17. CALCIUM-BINDING.
CC DOMAIN; 1155 1196 EGF-LIKE 18. CALCIUM-BINDING.
CC DOMAIN; 1197 1237 EGF-LIKE 19. CALCIUM-BINDING.
CC DOMAIN; 1238 1279 EGF-LIKE 20. CALCIUM-BINDING.

```

FT	DOMAIN	1280	1321	EGF-LIKE 21, CALCIUM-BINDING.	FT	DISULFID	1032	1044	BY SIMILARITY.
FT	DOMAIN	1322	1362	EGF-LIKE 22, CALCIUM-BINDING.	FT	DISULFID	1039	1053	BY SIMILARITY.
FT	DOMAIN	1363	1403	EGF-LIKE 23, CALCIUM-BINDING.	FT	DISULFID	1055	1068	BY SIMILARITY.
FT	DOMAIN	1404	1445	EGF-LIKE 24, CALCIUM-BINDING.	FT	DISULFID	1074	1086	BY SIMILARITY.
FT	DOMAIN	1446	1486	EGF-LIKE 25, CALCIUM-BINDING.	FT	DISULFID	1081	1095	BY SIMILARITY.
FT	DOMAIN	1487	1527	EGF-LIKE 26, CALCIUM-BINDING.	FT	DISULFID	1097	1111	BY SIMILARITY.
FT	REPEAT	1528	1605	TGFBP 4.	FT	DISULFID	1117	1129	BY SIMILARITY.
FT	DOMAIN	1606	1647	EGF-LIKE 27, CALCIUM-BINDING.	FT	DISULFID	1124	1138	BY SIMILARITY.
FT	DOMAIN	1648	1688	EGF-LIKE 28, CALCIUM-BINDING.	FT	DISULFID	1140	1153	BY SIMILARITY.
FT	REPEAT	1689	1765	TGFBP 5.	FT	DISULFID	1159	1171	BY SIMILARITY.
FT	DOMAIN	1766	1807	EGF-LIKE 29, CALCIUM-BINDING.	FT	DISULFID	1201	1212	BY SIMILARITY.
FT	DOMAIN	1808	1848	EGF-LIKE 30, CALCIUM-BINDING.	FT	DISULFID	1208	1221	BY SIMILARITY.
FT	DOMAIN	1849	1890	EGF-LIKE 31, CALCIUM-BINDING.	FT	DISULFID	1223	1236	BY SIMILARITY.
FT	DOMAIN	1891	1929	EGF-LIKE 32, CALCIUM-BINDING.	FT	DISULFID	1242	1254	BY SIMILARITY.
FT	DOMAIN	1930	1972	EGF-LIKE 33, CALCIUM-BINDING.	FT	DISULFID	1249	1263	BY SIMILARITY.
FT	DOMAIN	1973	2012	EGF-LIKE 34, CALCIUM-BINDING.	FT	DISULFID	1265	1278	BY SIMILARITY.
FT	DOMAIN	2013	2054	EGF-LIKE 35, CALCIUM-BINDING.	FT	DISULFID	1284	1296	BY SIMILARITY.
FT	REPEAT	2055	2126	TGFBP 6.	FT	DISULFID	1291	1305	BY SIMILARITY.
FT	DOMAIN	2127	2165	EGF-LIKE 36, CALCIUM-BINDING.	FT	DISULFID	1307	1320	BY SIMILARITY.
FT	DOMAIN	2166	2205	EGF-LIKE 37, CALCIUM-BINDING.	FT	DISULFID	1326	1339	BY SIMILARITY.
FT	DOMAIN	2206	2246	EGF-LIKE 38, CALCIUM-BINDING.	FT	DISULFID	1333	1348	BY SIMILARITY.
FT	DOMAIN	2247	2290	EGF-LIKE 39, CALCIUM-BINDING.	FT	DISULFID	1350	1361	BY SIMILARITY.
FT	DOMAIN	2291	2332	EGF-LIKE 40, CALCIUM-BINDING.	FT	DISULFID	1367	1380	BY SIMILARITY.
FT	REPEAT	2333	2401	TGFBP 7.	FT	DISULFID	1374	1389	BY SIMILARITY.
FT	DOMAIN	2402	2443	EGF-LIKE 41, CALCIUM-BINDING.	FT	DISULFID	1391	1402	BY SIMILARITY.
FT	DOMAIN	2444	2484	EGF-LIKE 42, CALCIUM-BINDING.	FT	DISULFID	1408	1420	BY SIMILARITY.
FT	DOMAIN	2485	2523	EGF-LIKE 43, CALCIUM-BINDING.	FT	DISULFID	1415	1429	BY SIMILARITY.
FT	DOMAIN	2524	2566	EGF-LIKE 44, CALCIUM-BINDING.	FT	DISULFID	1450	1461	BY SIMILARITY.
FT	DOMAIN	2567	2606	EGF-LIKE 45, CALCIUM-BINDING.	FT	DISULFID	1456	1470	BY SIMILARITY.
FT	DOMAIN	2607	2647	EGF-LIKE 46, CALCIUM-BINDING.	FT	DISULFID	1472	1485	BY SIMILARITY.
FT	DOMAIN	2648	2687	EGF-LIKE 47, CALCIUM-BINDING.	FT	DISULFID	1491	1502	BY SIMILARITY.
FT	DISULFID	85	94	BY SIMILARITY.	FT	DISULFID	1497	1511	BY SIMILARITY.
FT	DISULFID	89	100	BY SIMILARITY.	FT	DISULFID	1513	1526	BY SIMILARITY.
FT	DISULFID	102	111	BY SIMILARITY.	FT	DISULFID	1610	1622	BY SIMILARITY.
FT	DISULFID	119	129	BY SIMILARITY.	FT	DISULFID	1617	1631	BY SIMILARITY.
FT	DISULFID	123	134	BY SIMILARITY.	FT	DISULFID	1633	1646	BY SIMILARITY.
FT	DISULFID	136	145	BY SIMILARITY.	FT	DISULFID	1652	1663	BY SIMILARITY.
FT	DISULFID	150	160	BY SIMILARITY.	FT	DISULFID	1658	1672	BY SIMILARITY.
FT	DISULFID	154	166	BY SIMILARITY.	FT	DISULFID	1674	1687	BY SIMILARITY.
FT	DISULFID	168	177	BY SIMILARITY.	FT	DISULFID	1770	1782	BY SIMILARITY.
FT	DISULFID	250	262	BY SIMILARITY.	FT	DISULFID	1777	1791	BY SIMILARITY.
FT	DISULFID	257	271	BY SIMILARITY.	FT	DISULFID	1793	1806	BY SIMILARITY.
FT	DISULFID	273	286	BY SIMILARITY.	FT	DISULFID	1812	1824	BY SIMILARITY.
FT	DISULFID	292	304	BY SIMILARITY.	FT	DISULFID	1818	1833	BY SIMILARITY.
FT	DISULFID	299	313	BY SIMILARITY.	FT	DISULFID	1835	1847	BY SIMILARITY.
FT	DISULFID	313	328	BY SIMILARITY.	FT	DISULFID	1853	1865	BY SIMILARITY.
FT	DISULFID	453	465	BY SIMILARITY.	FT	DISULFID	1860	1874	BY SIMILARITY.
FT	DISULFID	460	474	BY SIMILARITY.	FT	DISULFID	1876	1889	BY SIMILARITY.
FT	DISULFID	476	488	BY SIMILARITY.	FT	DISULFID	1895	1905	BY SIMILARITY.
FT	DISULFID	494	504	BY SIMILARITY.	Query Match 12.1%; Score 678; DB 1; length 2877;				
FT	DISULFID	499	513	BY SIMILARITY.	Best Local Similarity 22.7%; Pred. No. 3,7e-37;				
FT	DISULFID	515	528	BY SIMILARITY.	Matches 270; Conservative 109; Mismatches 377; Indels 434; Gaps 51;				
FT	DISULFID	534	546	BY SIMILARITY.					
FT	DISULFID	541	555	BY SIMILARITY.					
FT	DISULFID	557	570	BY SIMILARITY.					
FT	DISULFID	576	587	BY SIMILARITY.					
FT	DISULFID	582	596	BY SIMILARITY.					
FT	DISULFID	598	611	BY SIMILARITY.					
FT	DISULFID	617	628	BY SIMILARITY.					
FT	DISULFID	623	637	BY SIMILARITY.					
FT	DISULFID	639	652	BY SIMILARITY.					
FT	DISULFID	727	739	BY SIMILARITY.					
FT	DISULFID	734	748	BY SIMILARITY.					
FT	DISULFID	750	763	BY SIMILARITY.					
FT	DISULFID	769	781	BY SIMILARITY.					
FT	DISULFID	776	790	BY SIMILARITY.					
FT	DISULFID	792	805	BY SIMILARITY.					
FT	DISULFID	811	821	BY SIMILARITY.					
FT	DISULFID	816	830	BY SIMILARITY.					
FT	DISULFID	832	845	BY SIMILARITY.					
FT	DISULFID	914	926	BY SIMILARITY.					
FT	DISULFID	921	935	BY SIMILARITY.					
FT	DISULFID	937	950	BY SIMILARITY.					



FT CHAIN 21 1394 LATENT TRANSFORMING GROWTH FACTOR BETA  
 FT DOMAIN 300 340 BINDING PROTEIN, ISOFORM 1S.  
 FT REPEAT 348 412 EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 546 587 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 588 629 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 630 670 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 671 710 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 711 751 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 752 792 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 793 833 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 834 874 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 875 916 EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 917 958 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 959 1001 EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).  
 FT REPEAT 1017 1084 REPEAT B.  
 FT DOMAIN 1097 1139 EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).  
 FT REPEAT 1190 1262 REPEAT C.  
 FT DOMAIN 1191 1262 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 1294 1334 EGF-LIKE 15.  
 FT DOMAIN 1335 1379 EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).  
 FT SITE 847 949 CELL ATTACHMENT SITE (POTENTIAL).  
 FT DISULFID 304 315 BY SIMILARITY.  
 FT DISULFID 310 324 BY SIMILARITY.  
 FT DISULFID 326 339 BY SIMILARITY.  
 FT DISULFID 550 562 BY SIMILARITY.  
 FT DISULFID 557 571 BY SIMILARITY.  
 FT DISULFID 573 586 BY SIMILARITY.  
 FT DISULFID 592 604 BY SIMILARITY.  
 FT DISULFID 599 613 BY SIMILARITY.  
 FT DISULFID 615 628 BY SIMILARITY.  
 FT DISULFID 634 645 BY SIMILARITY.  
 FT DISULFID 640 654 BY SIMILARITY.  
 FT DISULFID 657 669 BY SIMILARITY.  
 FT DISULFID 675 686 BY SIMILARITY.  
 FT DISULFID 681 695 BY SIMILARITY.  
 FT DISULFID 698 709 BY SIMILARITY.  
 FT DISULFID 715 726 BY SIMILARITY.  
 FT DISULFID 721 735 BY SIMILARITY.  
 FT DISULFID 737 750 BY SIMILARITY.  
 FT DISULFID 756 767 BY SIMILARITY.  
 FT DISULFID 762 776 BY SIMILARITY.  
 FT DISULFID 778 791 BY SIMILARITY.  
 FT DISULFID 797 808 BY SIMILARITY.  
 FT DISULFID 803 817 BY SIMILARITY.  
 FT DISULFID 819 832 BY SIMILARITY.  
 FT DISULFID 838 850 BY SIMILARITY.  
 FT DISULFID 845 859 BY SIMILARITY.  
 FT DISULFID 861 873 BY SIMILARITY.  
 FT DISULFID 879 891 BY SIMILARITY.  
 FT DISULFID 885 900 BY SIMILARITY.  
 FT DISULFID 902 915 BY SIMILARITY.  
 FT DISULFID 921 933 BY SIMILARITY.  
 FT DISULFID 927 942 BY SIMILARITY.  
 FT DISULFID 944 957 BY SIMILARITY.  
 FT DISULFID 963 975 BY SIMILARITY.  
 FT DISULFID 970 984 BY SIMILARITY.  
 FT DISULFID 986 1000 BY SIMILARITY.  
 FT DISULFID 1101 1114 BY SIMILARITY.  
 FT DISULFID 1109 1123 BY SIMILARITY.  
 FT DISULFID 1125 1138 BY SIMILARITY.  
 FT DISULFID 1144 1155 BY SIMILARITY.  
 FT DISULFID 1150 1164 BY SIMILARITY.  
 FT DISULFID 1166 1179 BY SIMILARITY.  
 FT DISULFID 1298 1309 BY SIMILARITY.  
 FT DISULFID 1304 1318 BY SIMILARITY.  
 FT DISULFID 1320 1333 BY SIMILARITY.  
 FT DISULFID 1339 1354 BY SIMILARITY.  
 FT DISULFID 1349 1363 BY SIMILARITY.  
 FT DISULFID 1365 1378 BY SIMILARITY.  
 FT MOD\_RES 647 647 HYDROXYLATION.  
 FT MOD\_RES 810 810 HYDROXYLATION.  
 FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 870 870 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 923 923 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT FTID-CAR 000184 /FTID-CAR 000184  
 FT CONFLICT 365 365 Y -> H (IN REF. 2).  
 FT CONFLICT 504 504 A -> T (IN REF. 2).  
 FT CONFLICT 1334 1334 F -> V (IN REF. 2).  
 SC SEQUENCE 1394 AA: 152791 MW: 27581.440327C1 CRC64;  
 Query Match 11.8; Score 624.5; DB 1; Length 1394;  
 Best Local Similarity 24.9%; Pred. No. 66-34;  
 Matches 24; Conservative 94; Mismatches 334; Indels 297; Gaps 48;  
 QY 45 DVDECAQGLDDCHADALCALNTPTSYKSCPKGVQ--GEGROCEDIDECNELNGSCVHD- 101  
 DB 546 EINECTWNPICGA-GHCINLPVRYTCICVEGYRSEQQKCVDIDEC-TQVCHLCSQGR 603  
 QY 102 CLNIPGNYRCTPDGPMJAHDHNCILDVDECLNNGSGCHTCTNYMGSEYEC-CKEKGFLP 160  
 DB 604 CENTEGSFELICPAGPMASEEGTNCIDVDJCLRPDVCYGECHGCVTVGAFCRCYCDGVM 663  
 QY 161 SDNQHTCIHRESESLSCMNKDHGCHICKAEPGSGVACRPGFELAKNQRDCILTCNHG 220  
 DB 664 TQR-----GRCEDIDELNPSTCPDEQVNSP-GSYQC-----VPCFEG 701  
 QY 221 NGCOHSCDUTADGPESCSCHPQYKMTDGRSCLEDECTVLEVTESNTTSVWZGDKEVRR 280  
 DB 702 FRGNQCCLDVE-----CUE----- 717  
 QY 281 LLMETCAVNGSCDRCKDTSTGVHSCVPVGTLLDQKCKKIDIDECQTRNGGCHFFCKY 340  
 DB 718 --PNVCA--NG-----DCSNLESGYMSCHKGVTYTPDKHCHREDIDECQGNLVNGCKN 769  
 QY 341 IVGSFQCGKCKFKLLTCKSKQVDECSUDRTCH-SCIRHEGTFAACNRYGTYLGF- 398  
 DB 770 TEGSFCTCGQGVLSAAKQCEIDCEOKHLCAGGCRNTEGSCFQVCDQGYASGAG 829  
 QY 399 THCGDTNECSINNGGCGQ--VCVNTVSGSEYEQCHPGYKJHWNK-----KDCVEVKGLPTS 452  
 DB 830 CHCEDINELEDKSVQQRGDCNTAGSYDCTCPDGFQLDNKTCCQINEC-EHPGJ----- 584  
 QY 453 VSPRVSJHCKSG--GGDGCF-LRCHSGIHLSSVTTIRTSTVTKJREGKSLXNAELF 508  
 DB 885 -----GPOGECUNTFGSHCVCCQGSFSAQRT-----CRDI 918  
 QY 509 PEGLRPALPEKH---SSVKESFRYINLTCSGKQVF-----GAPGRP-- 547  
 DB 919 DECVNNTVCDSHGFCDNTAGSFRCL--CYQGFQAFQDQGGCVNVEGELLGVCGEAFC 975  
 QY 549 STPKEMFIV-----EPELETKOKEVTASCDLSCTVKTREKLRKA-NTLRKA----- 595  
 DB 976 ENVEGSFLVCADENQSEYSPYTGQCRSRSTSLDQVDVDQFKEEKCECYNLNDASLCNV 1035  
 QY 596 ---VHREQFHLCL---GMNLDVAKKPPRTSEROAESCQVGGH-----AEN- 636  
 DB 1036 LAENVTKQECCTSGAGWGDNCEIFPCPVLTGTAETEMCPKCKGFPVPAESSEAGENY 1095  
 QY 637 ---QCV-----SCRAGTYVDGARERC-----LCPNGTF 662  
 DB 1096 KDADECLLFQOEICKNGFCNLTNPGEYCYCKQGYVDPVKLQCFDMECDQSPSSCIDGQC 1155  
 QY 663 QNEEGQWTC-----EPCRPNGS-----G 681  
 DB 1156 VNTGSGNCFCTHPVLDASEKRCIRPASENQCIBETDVQDLCHWHLSDVVCGRPLWG 1215  
 QY 682 ALXT-----PEAWNMSECGGLCPGGEYSADGAPCOLCALGTGFQPEAGRTSCF----- 729  
 DB 1216 KQTTYTECCCLYGEAWGM-QC-ALC-PLKDSDDYACLNIPIVTRRQP-YGRDALVDSEF 1271  
 QY 730 ---PCGGGLATKHQGATSFQDCETRVQCSGPHGYNTTTHRCIRCPVG-----TYQPEF 779





Query Match 11.1%; Score 624.5; DB 1; Length 1595;  
 Best Local Similarity 24.9%; Pred. No. 7e-34;  
 Matches 240; Conservative 94; Mismatches 334; Indels 297; Gaps 48;

QY 45 DVDECAQGLDDHADALQNTTYSKCSKPGYQ--GEGRCQEDIDECNENLNGCVHD- 101  
 DB 747 EINECTVNPDIQGA-GHCINLPVRYTCIEGYRFEQKQVD-DEC-TQVH-CSSQGR 304

QY 102 CLNIPGNYRCTCDGFMALDHGNCIDVDECLNNGGCOHTCVNMGSECC-CCKEGSFL 160  
 DB 805 CENTGFSFICPAGFMASEEGTNCIDVDECLRPDVCGEHCWNTVGAFCYCDGSRM 964

QY 161 SONQHTCIHRSEGLSCMNKDHGSHICKAEPGRSVACRCRQFELAKQORCIITCNKG 220  
 DB 865 TOR-----GRCED:DECLNPSTCPDQCVNSP-GSYQC-----VPTGEG 902

QY 221 NGCQHSDDTADGPCCSCHPQVYMTDGRSCLEREDTVLENTSTVVDGKVKRR 280  
 DB 903 FROWNGQCLDVDE-----CUE----- 918

QY 281 LLMETCAVNGGCDRTCKDTSTGVHSCVPGFTLQLDGXTCKIDECOC-RNSGCHFCNK 340  
 DB 919 --PNTCA--NG-----DCSNLEGS:WCSCHKGYTRTPDHKHCRIIDECQQGNLCVNGQCN 970

QY 341 IVGSFDCGCKGFKLLTDEKSCQDVDECSLDRTCDH-SCINHPGTACACNRGYLYGF- 398  
 DB 971 TEGSFRTCTCGYOLSAKQDQEDIDECQHRHLCAHQCRNTEGSGFCQCDGYRASGLG 1030

QY 399 THCGDPTNCSINNGGCOQ-VCVNTVGSYBCQHPGYKLHWK-----KDCVEKGLLPTS 452  
 DB 1031 DHCEDINECLDKSVQCRGDCINTAGSYDCTPDGFLQDNDKTCQDINBC-EHPGJL----- 1085

QY 453 VSPRVLHCKSG--GGQGCFLRCHSGTHLSSDVTTITRTSVFKLNEGKSLKNAELF 508  
 DB 1086 -----CQPGQEC:NTGESHFCVCCQGFSTISADGT-----CEDI 1119

QY 509 PEGLRPALPEKH---SSVKESPRYVNLTCSSGKQVP-----GAPGRP-- 547  
 DB 1:20 DECNNNTVCDSHGFCNDTAGSPFCT---CYQGFAFCQDQGVVNECELLSSVGEAF 1176

QY 548 STPKEMFITY-----EPELETNQKEVTSACLSLCIVKTEKRLKAIITLKA----- 595

Db 1177 ENVEGSFLVCADENQVEYSPMTCCQBSRTSDLDVDVDQPKKEKKECYNYLNDASLCNV 1236  
 QY 596 ----VHREQPHLQLS---GMNLDVAKKPPRTSERQAESCGVQGH-----AEN- 636  
 Db 1237 LAPNVTQKECCCTSGAGWGNCEIFPCPVLGTAETEMCPKGFVPAGSESSAAGENY 1296  
 QY 637 ----QCV-----SCRACTVYDGAERCI-----LCPNGTF 662  
 Db 1297 KDADECLFQGEIKCKNGFCINTRPGVECYCKQCTYYDPVKLOCFMDECDQSSCIDGQC 1356  
 QY 663 QNEEGQWTC-----EPCPRPGNS-----G 681  
 Db 1357 VNTGSSVNCFTHPMVLDASEKRCIRPAESNBOIBETDVYQDLWEHLSDVYVCSRPLVG 1416  
 QY 682 ALKT-----PEAWNMBEGGLQCPGEYSADGFAPOCALGTQPPAGRTSCF---- 729  
 Db 1417 KQTTYTECCCLYGEAWGM-QC-ALC-PLKDSDDYAOCLNIPVTGRRQP-YGRDALVDFSE 1472  
 QY 730 ---PCGGGLATKHOGATSFQDCETRVQCSGPHFYNTTTHRCIRCPVG-----TYQPEF 779  
 Db 1473 QYTFEADPFYIQDFLNSFEELQAE-SC--GLINGCENGRCVRVQEGYTCDCLDGYHLST 1529  
 QY 780 GKNNVSCPGNTTTFDGSNTN-ITCKNRRCQGEIGDF-----TGYES--PNY--PGNY 829  
 Db 1530 AKMTCPD-----VNECDELNNRMSLCKNAK:NTDGSYKCLCLPGVYVPSDKPNYCTPLNT 1584  
 QY 830 PANTE 834  
 Db 1585 ALNJE 1599

RESULT 11  
 ID: LBS\_VXUSE STANDARD: PRT: 1359 AA.  
 AC Q8CG18; Q8BNW7; Q8C7F5; Q8C1RC;  
 DT 15-SEP-2003 (Rel. 42, Created:  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Latent transforming growth factor beta binding protein, isoform 1S  
 DE precursor (LTBP-1) (transforming growth factor beta-1 binding protein  
 DE 1) (TGF-beta1-SP-1).  
 GN LTBP1.  
 CS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN 1.  
 RP SEQUENCE FROM N.A.  
 RX PubMed=12711388;  
 RA Noguera I., Obata H., Gualandris A., Cowin P., Rifkin D.B.;  
 RT "Molecular cloning of the mouse Ltbp-1 gene reveals tissue specific  
 expression of alternatively spliced forms.";  
 RL Gene 308:31-41(2003).  
 RN 2.  
 RP SEQUENCE FROM N.A.  
 RX PubMed=12711389;  
 RA Weiskirchen R., Moser M., Guenther K., Weiskirchen S., Gressner A.M.;  
 RT "The murine latent transforming growth factor-beta binding protein  
 (Ltbp-1) is alternatively spliced, and maps to a region syntenic to  
 human chromosome 2p21-22.";  
 RL Gene 308:43-52(2003).  
 RN 3.  
 RP SEQUENCE OF 788-1389 FROM N.A.  
 PC STRAIN=C57BL/6J; TISSUE=Aorta, Liver, and Vein;  
 RX PubMed=12466851;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nakaide T., Osato N., Saito R., Suzuki H., Yamahata T., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schrim D.M., Knapin A., Matsuda H., Baralov S., Beise K.W.,  
 RA Sake C.A., Bratt D., Brusik V., Chothia C., Corbani J.E., Cousins S.,  
 RA DeLa E., Dragani Z.A., Fletcher C.F., Forrest A., Frazer K.S.,

RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,  
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
RA Maglott D.R., Malcais L., Marchionni L., McKenzie L., Miki H.,  
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
RA Sultana R., Tagner L., Taylor M.S., Teasdale R.D., Tomita M.,  
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Korno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,  
RA Yasuishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.,  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs";  
RL Nature 420:563-573(2002).  
CC - SUBMIT: The large latent complex of TGF-beta1 from platelets is  
CC composed of the TGF-beta1 molecule noncovalently associated with a  
CC disulfide-bonded complex of a dimer of the N-terminal propeptide  
CC of the TGF-beta1 precursor and a third component denoted TGF-  
CC beta1-BP. TGF-beta1-BP does not bind directly to active TGF-BETA1.  
CC Binds to fibrillin (By similarity).  
CC - SUBCELLULAR LOCATION: Secreted (By similarity).  
CC - ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=IS;  
CC IsoId=Q8CG18-1; Sequence=Displayed;  
CC Name=IL;  
CC IsoId=Q8CG19-1; Sequence=External;  
CC - PTM: Contains hydroxylated asparagine residues (By similarity).  
CC - PTM: The N-terminus is blocked (By similarity).  
CC - SIMILARITY: Contains 16 EGF-like domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF346465; AAN77251.1; JOINED.  
DR EMBL; AF346438; AAN77251.1; JOINED.  
DR EMBL; AF346439; AAN77251.1; JOINED.  
DR EMBL; AF346440; AAN77251.1; JOINED.  
DR EMBL; AF346441; AAN77251.1; JOINED.  
DR EMBL; AF346442; AAN77251.1; JOINED.  
DR EMBL; AF346443; AAN77251.1; JOINED.  
DR EMBL; AF346444; AAN77251.1; JOINED.  
DR EMBL; AF346445; AAN77251.1; JOINED.  
DR EMBL; AF346446; AAN77251.1; JOINED.  
DR EMBL; AF346447; AAN77251.1; JOINED.  
DR EMBL; AF346448; AAN77251.1; JOINED.  
DR EMBL; AF346449; AAN77251.1; JOINED.  
DR EMBL; AF346450; AAN77251.1; JOINED.  
DR EMBL; AF346451; AAN77251.1; JOINED.  
DR EMBL; AF346452; AAN77251.1; JOINED.  
DR EMBL; AF346453; AAN77251.1; JOINED.  
DR EMBL; AF346454; AAN77251.1; JOINED.  
DR EMBL; AF346455; AAN77251.1; JOINED.  
DR EMBL; AF346456; AAN77251.1; JOINED.  
DR EMBL; AF346457; AAN77251.1; JOINED.  
DR EMBL; AF346458; AAN77251.1; JOINED.  
DR EMBL; AF346459; AAN77251.1; JOINED.  
DR EMBL; AF346460; AAN77251.1; JOINED.  
DR EMBL; AF346461; AAN77251.1; JOINED.  
DR EMBL; AF346462; AAN77251.1; JOINED.  
DR EMBL; AF346463; AAN77251.1; JOINED.  
DR EMBL; AF346464; AAN77251.1; JOINED.  
DR EMBL; AY431161; AAN38831.1; ALT\_SEQ.  
DR EMBL; AK50390; BAC34222.1; .  
DR EMBL; AK380024; BAC37809.1; .  
DR XGD; MGI:109151; Lbbp.  
DR InterPro; IPRO00152; Asx\_hydroxy.  
DR InterPro; IPRO01881; EGF\_Ca.  
DR InterPro; IPRO036209; EGF\_Like.  
DR InterPro; IPRO02212; Fibril-assoc.  
DR InterPro; IPRO06210; EGF.  
DR Pfam; PF00228; EGF; 15.  
DR Pfam; PF00685; TS; 4.  
DR SMART; SM00181; EGF; 17.  
DR SMART; SM00179; EGF\_CA; 16.  
DR PROSITE; PS0010; ASX\_HYDROXY; 13.  
DR PROSITE; PS0022; EGF; 1.  
DR PROSITE; PS01186; EGF; 2; 11.  
DR PROSITE; PS01187; EGF\_CA; 15.  
DR Growth factor binding; Repeat; EGF-like domain;  
KW Hydroxylation; Signal; Glycoprotein;  
KW Alternative splicing.  
FT SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 1389 LATENT TRANSFORMING GROWTH FACTOR BETA  
FT BINDING PROTEIN, ISOFORM 1S.  
FT EGF-LIKE 1, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT A.  
FT DOMAIN 295 335 EGF-LIKE 2, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT B.  
FT DOMAIN 343 408 EGF-LIKE 3, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT C.  
FT DOMAIN 542 583 EGF-LIKE 4, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT D.  
FT DOMAIN 584 625 EGF-LIKE 5, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT E.  
FT DOMAIN 626 666 EGF-LIKE 6, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT F.  
FT DOMAIN 667 706 EGF-LIKE 7, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT G.  
FT DOMAIN 707 747 EGF-LIKE 8, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT H.  
FT DOMAIN 748 788 EGF-LIKE 9, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT I.  
FT DOMAIN 789 829 EGF-LIKE 10, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT J.  
FT DOMAIN 830 870 EGF-LIKE 11, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT K.  
FT DOMAIN 871 912 EGF-LIKE 12, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT L.  
FT DOMAIN 913 954 EGF-LIKE 13, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT M.  
FT DOMAIN 955 997 EGF-LIKE 14, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT N.  
FT DOMAIN 1012 1079 EGF-LIKE 15, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT O.  
FT DOMAIN 1032 1134 EGF-LIKE 16, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT P.  
FT DOMAIN 1135 1175 EGF-LIKE 17, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT Q.  
FT DOMAIN 1185 1257 EGF-LIKE 18, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT R.  
FT DOMAIN 1269 1330 EGF-LIKE 19, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT S.  
FT DOMAIN 1330 1374 EGF-LIKE 20, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT T.  
FT DOMAIN 1374 1414 EGF-LIKE 21, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT U.  
FT DOMAIN 1414 1454 EGF-LIKE 22, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT V.  
FT DOMAIN 1454 1494 EGF-LIKE 23, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT W.  
FT DOMAIN 1494 1534 EGF-LIKE 24, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT X.  
FT DOMAIN 1534 1574 EGF-LIKE 25, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT Y.  
FT DOMAIN 1574 1614 EGF-LIKE 26, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT Z.  
FT DOMAIN 1614 1654 EGF-LIKE 27, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AA.  
FT DOMAIN 1654 1694 EGF-LIKE 28, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AB.  
FT DOMAIN 1694 1734 EGF-LIKE 29, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AC.  
FT DOMAIN 1734 1774 EGF-LIKE 30, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AD.  
FT DOMAIN 1774 1814 EGF-LIKE 31, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AE.  
FT DOMAIN 1814 1854 EGF-LIKE 32, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AF.  
FT DOMAIN 1854 1894 EGF-LIKE 33, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AG.  
FT DOMAIN 1894 1934 EGF-LIKE 34, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AH.  
FT DOMAIN 1934 1974 EGF-LIKE 35, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AI.  
FT DOMAIN 1974 2014 EGF-LIKE 36, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AJ.  
FT DOMAIN 2014 2054 EGF-LIKE 37, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AK.  
FT DOMAIN 2054 2094 EGF-LIKE 38, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AL.  
FT DOMAIN 2094 2134 EGF-LIKE 39, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AM.  
FT DOMAIN 2134 2174 EGF-LIKE 40, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AN.  
FT DOMAIN 2174 2214 EGF-LIKE 41, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AO.  
FT DOMAIN 2214 2254 EGF-LIKE 42, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AP.  
FT DOMAIN 2254 2294 EGF-LIKE 43, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AQ.  
FT DOMAIN 2294 2334 EGF-LIKE 44, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AR.  
FT DOMAIN 2334 2374 EGF-LIKE 45, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AS.  
FT DOMAIN 2374 2414 EGF-LIKE 46, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AT.  
FT DOMAIN 2414 2454 EGF-LIKE 47, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AU.  
FT DOMAIN 2454 2494 EGF-LIKE 48, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AV.  
FT DOMAIN 2494 2534 EGF-LIKE 49, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AW.  
FT DOMAIN 2534 2574 EGF-LIKE 50, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AX.  
FT DOMAIN 2574 2614 EGF-LIKE 51, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AY.  
FT DOMAIN 2614 2654 EGF-LIKE 52, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT AZ.  
FT DOMAIN 2654 2694 EGF-LIKE 53, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BA.  
FT DOMAIN 2694 2734 EGF-LIKE 54, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BB.  
FT DOMAIN 2734 2774 EGF-LIKE 55, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BC.  
FT DOMAIN 2774 2814 EGF-LIKE 56, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BD.  
FT DOMAIN 2814 2854 EGF-LIKE 57, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BE.  
FT DOMAIN 2854 2894 EGF-LIKE 58, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BF.  
FT DOMAIN 2894 2934 EGF-LIKE 59, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BG.  
FT DOMAIN 2934 2974 EGF-LIKE 60, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BH.  
FT DOMAIN 2974 3014 EGF-LIKE 61, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BI.  
FT DOMAIN 3014 3054 EGF-LIKE 62, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BJ.  
FT DOMAIN 3054 3094 EGF-LIKE 63, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BK.  
FT DOMAIN 3094 3134 EGF-LIKE 64, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BL.  
FT DOMAIN 3134 3174 EGF-LIKE 65, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BM.  
FT DOMAIN 3174 3214 EGF-LIKE 66, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BN.  
FT DOMAIN 3214 3254 EGF-LIKE 67, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BO.  
FT DOMAIN 3254 3294 EGF-LIKE 68, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BP.  
FT DOMAIN 3294 3334 EGF-LIKE 69, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BQ.  
FT DOMAIN 3334 3374 EGF-LIKE 70, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BR.  
FT DOMAIN 3374 3414 EGF-LIKE 71, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BS.  
FT DOMAIN 3414 3454 EGF-LIKE 72, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BT.  
FT DOMAIN 3454 3494 EGF-LIKE 73, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BU.  
FT DOMAIN 3494 3534 EGF-LIKE 74, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BV.  
FT DOMAIN 3534 3574 EGF-LIKE 75, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BW.  
FT DOMAIN 3574 3614 EGF-LIKE 76, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BX.  
FT DOMAIN 3614 3654 EGF-LIKE 77, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BY.  
FT DOMAIN 3654 3694 EGF-LIKE 78, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT BZ.  
FT DOMAIN 3694 3734 EGF-LIKE 79, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CA.  
FT DOMAIN 3734 3774 EGF-LIKE 80, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CB.  
FT DOMAIN 3774 3814 EGF-LIKE 81, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CC.  
FT DOMAIN 3814 3854 EGF-LIKE 82, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CD.  
FT DOMAIN 3854 3894 EGF-LIKE 83, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CE.  
FT DOMAIN 3894 3934 EGF-LIKE 84, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CF.  
FT DOMAIN 3934 3974 EGF-LIKE 85, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CG.  
FT DOMAIN 3974 4014 EGF-LIKE 86, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CH.  
FT DOMAIN 4014 4054 EGF-LIKE 87, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CI.  
FT DOMAIN 4054 4094 EGF-LIKE 88, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CJ.  
FT DOMAIN 4094 4134 EGF-LIKE 89, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CK.  
FT DOMAIN 4134 4174 EGF-LIKE 90, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CL.  
FT DOMAIN 4174 4214 EGF-LIKE 91, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CM.  
FT DOMAIN 4214 4254 EGF-LIKE 92, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CN.  
FT DOMAIN 4254 4294 EGF-LIKE 93, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CO.  
FT DOMAIN 4294 4334 EGF-LIKE 94, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CP.  
FT DOMAIN 4334 4374 EGF-LIKE 95, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CQ.  
FT DOMAIN 4374 4414 EGF-LIKE 96, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CR.  
FT DOMAIN 4414 4454 EGF-LIKE 97, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CS.  
FT DOMAIN 4454 4494 EGF-LIKE 98, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CT.  
FT DOMAIN 4494 4534 EGF-LIKE 99, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CU.  
FT DOMAIN 4534 4574 EGF-LIKE 100, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CV.  
FT DOMAIN 4574 4614 EGF-LIKE 101, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CW.  
FT DOMAIN 4614 4654 EGF-LIKE 102, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CX.  
FT DOMAIN 4654 4694 EGF-LIKE 103, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CY.  
FT DOMAIN 4694 4734 EGF-LIKE 104, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT CZ.  
FT DOMAIN 4734 4774 EGF-LIKE 105, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DA.  
FT DOMAIN 4774 4814 EGF-LIKE 106, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DB.  
FT DOMAIN 4814 4854 EGF-LIKE 107, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DC.  
FT DOMAIN 4854 4894 EGF-LIKE 108, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DD.  
FT DOMAIN 4894 4934 EGF-LIKE 109, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DE.  
FT DOMAIN 4934 4974 EGF-LIKE 110, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DF.  
FT DOMAIN 4974 5014 EGF-LIKE 111, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DG.  
FT DOMAIN 5014 5054 EGF-LIKE 112, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DH.  
FT DOMAIN 5054 5094 EGF-LIKE 113, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DI.  
FT DOMAIN 5094 5134 EGF-LIKE 114, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DJ.  
FT DOMAIN 5134 5174 EGF-LIKE 115, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DK.  
FT DOMAIN 5174 5214 EGF-LIKE 116, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DL.  
FT DOMAIN 5214 5254 EGF-LIKE 117, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DM.  
FT DOMAIN 5254 5294 EGF-LIKE 118, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DN.  
FT DOMAIN 5294 5334 EGF-LIKE 119, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DO.  
FT DOMAIN 5334 5374 EGF-LIKE 120, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DP.  
FT DOMAIN 5374 5414 EGF-LIKE 121, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DQ.  
FT DOMAIN 5414 5454 EGF-LIKE 122, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DR.  
FT DOMAIN 5454 5494 EGF-LIKE 123, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DS.  
FT DOMAIN 5494 5534 EGF-LIKE 124, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DT.  
FT DOMAIN 5534 5574 EGF-LIKE 125, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DU.  
FT DOMAIN 5574 5614 EGF-LIKE 126, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DV.  
FT DOMAIN 5614 5654 EGF-LIKE 127, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DW.  
FT DOMAIN 5654 5694 EGF-LIKE 128, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DX.  
FT DOMAIN 5694 5734 EGF-LIKE 129, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DY.  
FT DOMAIN 5734 5774 EGF-LIKE 130, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT DZ.  
FT DOMAIN 5774 5814 EGF-LIKE 131, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EA.  
FT DOMAIN 5814 5854 EGF-LIKE 132, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EB.  
FT DOMAIN 5854 5894 EGF-LIKE 133, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EC.  
FT DOMAIN 5894 5934 EGF-LIKE 134, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT ED.  
FT DOMAIN 5934 5974 EGF-LIKE 135, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EE.  
FT DOMAIN 5974 6014 EGF-LIKE 136, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EF.  
FT DOMAIN 6014 6054 EGF-LIKE 137, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EG.  
FT DOMAIN 6054 6094 EGF-LIKE 138, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EH.  
FT DOMAIN 6094 6134 EGF-LIKE 139, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EI.  
FT DOMAIN 6134 6174 EGF-LIKE 140, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EJ.  
FT DOMAIN 6174 6214 EGF-LIKE 141, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EK.  
FT DOMAIN 6214 6254 EGF-LIKE 142, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EL.  
FT DOMAIN 6254 6294 EGF-LIKE 143, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EM.  
FT DOMAIN 6294 6334 EGF-LIKE 144, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EN.  
FT DOMAIN 6334 6374 EGF-LIKE 145, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EO.  
FT DOMAIN 6374 6414 EGF-LIKE 146, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EP.  
FT DOMAIN 6414 6454 EGF-LIKE 147, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EQ.  
FT DOMAIN 6454 6494 EGF-LIKE 148, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT ER.  
FT DOMAIN 6494 6534 EGF-LIKE 149, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT ES.  
FT DOMAIN 6534 6574 EGF-LIKE 150, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT ET.  
FT DOMAIN 6574 6614 EGF-LIKE 151, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EU.  
FT DOMAIN 6614 6654 EGF-LIKE 152, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EV.  
FT DOMAIN 6654 6694 EGF-LIKE 153, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EW.  
FT DOMAIN 6694 6734 EGF-LIKE 154, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EX.  
FT DOMAIN 6734 6774 EGF-LIKE 155, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EY.  
FT DOMAIN 6774 6814 EGF-LIKE 156, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT EZ.  
FT DOMAIN 6814 6854 EGF-LIKE 157, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FA.  
FT DOMAIN 6854 6894 EGF-LIKE 158, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FB.  
FT DOMAIN 6894 6934 EGF-LIKE 159, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FC.  
FT DOMAIN 6934 6974 EGF-LIKE 160, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FD.  
FT DOMAIN 6974 7014 EGF-LIKE 161, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FE.  
FT DOMAIN 7014 7054 EGF-LIKE 162, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FF.  
FT DOMAIN 7054 7094 EGF-LIKE 163, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FG.  
FT DOMAIN 7094 7134 EGF-LIKE 164, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FH.  
FT DOMAIN 7134 7174 EGF-LIKE 165, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FI.  
FT DOMAIN 7174 7214 EGF-LIKE 166, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FJ.  
FT DOMAIN 7214 7254 EGF-LIKE 167, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FK.  
FT DOMAIN 7254 7294 EGF-LIKE 168, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FL.  
FT DOMAIN 7294 7334 EGF-LIKE 169, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FM.  
FT DOMAIN 7334 7374 EGF-LIKE 170, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FN.  
FT DOMAIN 7374 7414 EGF-LIKE 171, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FO.  
FT DOMAIN 7414 7454 EGF-LIKE 172, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FP.  
FT DOMAIN 7454 7494 EGF-LIKE 173, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FQ.  
FT DOMAIN 7494 7534 EGF-LIKE 174, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FR.  
FT DOMAIN 7534 7574 EGF-LIKE 175, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FS.  
FT DOMAIN 7574 7614 EGF-LIKE 176, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FT.  
FT DOMAIN 7614 7654 EGF-LIKE 177, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FU.  
FT DOMAIN 7654 7694 EGF-LIKE 178, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FV.  
FT DOMAIN 7694 7734 EGF-LIKE 179, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FW.  
FT DOMAIN 7734 7774 EGF-LIKE 180, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FX.  
FT DOMAIN 7774 7814 EGF-LIKE 181, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FY.  
FT DOMAIN 7814 7854 EGF-LIKE 182, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT FZ.  
FT DOMAIN 7854 7894 EGF-LIKE 183, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GA.  
FT DOMAIN 7894 7934 EGF-LIKE 184, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GB.  
FT DOMAIN 7934 7974 EGF-LIKE 185, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GC.  
FT DOMAIN 7974 8014 EGF-LIKE 186, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GD.  
FT DOMAIN 8014 8054 EGF-LIKE 187, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GE.  
FT DOMAIN 8054 8094 EGF-LIKE 188, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GF.  
FT DOMAIN 8094 8134 EGF-LIKE 189, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GH.  
FT DOMAIN 8134 8174 EGF-LIKE 190, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GI.  
FT DOMAIN 8174 8214 EGF-LIKE 191, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GJ.  
FT DOMAIN 8214 8254 EGF-LIKE 192, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GK.  
FT DOMAIN 8254 8294 EGF-LIKE 193, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GL.  
FT DOMAIN 8294 8334 EGF-LIKE 194, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GM.  
FT DOMAIN 8334 8374 EGF-LIKE 195, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GN.  
FT DOMAIN 8374 8414 EGF-LIKE 196, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GO.  
FT DOMAIN 8414 8454 EGF-LIKE 197, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GP.  
FT DOMAIN 8454 8494 EGF-LIKE 198, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GQ.  
FT DOMAIN 8494 8534 EGF-LIKE 199, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GR.  
FT DOMAIN 8534 8574 EGF-LIKE 200, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GS.  
FT DOMAIN 8574 8614 EGF-LIKE 201, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GT.  
FT DOMAIN 8614 8654 EGF-LIKE 202, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GU.  
FT DOMAIN 8654 8694 EGF-LIKE 203, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GV.  
FT DOMAIN 8694 8734 EGF-LIKE 204, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GW.  
FT DOMAIN 8734 8774 EGF-LIKE 205, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GX.  
FT DOMAIN 8774 8814 EGF-LIKE 206, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GY.  
FT DOMAIN 8814 8854 EGF-LIKE 207, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT GZ.  
FT DOMAIN 8854 8894 EGF-LIKE 208, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HA.  
FT DOMAIN 8894 8934 EGF-LIKE 209, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HB.  
FT DOMAIN 8934 8974 EGF-LIKE 210, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HC.  
FT DOMAIN 8974 9014 EGF-LIKE 211, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HD.  
FT DOMAIN 9014 9054 EGF-LIKE 212, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HE.  
FT DOMAIN 9054 9094 EGF-LIKE 213, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HF.  
FT DOMAIN 9094 9134 EGF-LIKE 214, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HG.  
FT DOMAIN 9134 9174 EGF-LIKE 215, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HH.  
FT DOMAIN 9174 9214 EGF-LIKE 216, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HI.  
FT DOMAIN 9214 9254 EGF-LIKE 217, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HJ.  
FT DOMAIN 9254 9294 EGF-LIKE 218, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HK.  
FT DOMAIN 9294 9334 EGF-LIKE 219, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HL.  
FT DOMAIN 9334 9374 EGF-LIKE 220, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HM.  
FT DOMAIN 9374 9414 EGF-LIKE 221, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HN.  
FT DOMAIN 9414 9454 EGF-LIKE 222, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HO.  
FT DOMAIN 9454 9494 EGF-LIKE 223, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HP.  
FT DOMAIN 9494 9534 EGF-LIKE 224, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HQ.  
FT DOMAIN 9534 9574 EGF-LIKE 225, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HR.  
FT DOMAIN 9574 9614 EGF-LIKE 226, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HS.  
FT DOMAIN 9614 9654 EGF-LIKE 227, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HT.  
FT DOMAIN 9654 9694 EGF-LIKE 228, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HU.  
FT DOMAIN 9694 9734 EGF-LIKE 229, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HV.  
FT DOMAIN 9734 9774 EGF-LIKE 230, CALCIUM-BINDING  
(POTENTIAL).  
FT REPEAT HW.  
FT DOMAIN 9774 9814 EGF-LIKE 2

```
FT DISULFID 694 705 POTENTIAL.
FT DISULFID 711 722 POTENTIAL.
FT DISULFID 717 731 POTENTIAL.
FT DISULFID 733 746 POTENTIAL.
FT DISULFID 752 763 POTENTIAL.
FT DISULFID 758 772 POTENTIAL.
FT DISULFID 774 787 POTENTIAL.
FT DISULFID 793 804 POTENTIAL.
FT DISULFID 799 813 POTENTIAL.
FT DISULFID 815 828 POTENTIAL.
FT DISULFID 834 846 POTENTIAL.
FT DISULFID 841 855 POTENTIAL.
FT DISULFID 857 869 POTENTIAL.
FT DISULFID 875 887 POTENTIAL.

Query Match 10.7%; Score 601; DB 1; Length 1389;
Best Local Similarity 24.2%; Pred. No. 2.2e-12;
Matches 230; Conservative 105; Mismatches 305; Indels 310; Gaps 52;

QY 45 DVDECAQGLDECHADALCQNTPTSYKCSCKPGYQG--EGROCEDIDECQNE--LNGGCV 99
Db 584 DIDECAQVRLHC--SQRCENTGESFLVCVPAGMASEEGTNCIDVDECLRPDMCRDGRCI 642
QY 100 H-----DCLNIPGNRYCT--CFDGFMLAH 121
Db 643 NTAGAFRCYCDSGYRMRRGYCEDIDELKPTCEEQCVNTPGSYQCVPTCEGFR--- 699
QY 122 DGHN--CLYVDECLENNGGQCHTCVNVGSEYCCCKEGFFLSDNQHTC--IHRSEGLSC 177
Db 700 -GNGQCILDVDECLQPKVCTNGSCTNLEGSYMCSCHRGYSPTPEHRHCCDIDECQGNLC 759
QY 178 MNKHGDCSHICKAPRGSVACERPGFEL--AKVQRDCILTNHNGGQCH--SCDGTADG 234
Db 759 MNGQ-----CRNTD-GSFRCTCGQVQLSAKDQCEDIDECSEH-HHLCSHQCRNTEGS 810
QY 235 PEGSCRPQYKMTJGRCCLREBTIVLEVTESTNTSVVQSGKRVKRRLLMETCAVNGGCC 294
Db 911 FQVCVQGYRASVLGDHC--ED-----INE-----CLEDSSVCQG-- 844
QY 295 RTCKDSTGVHSCPVGFTQLQDKTKCIDECOTRNGGCDHF--CKNTVGSFDCGCKKG 352
Db 845 -DCINTAGSDTCTCPDGFQNL-DNKGCDLINE-AQGLCGSHGECNTQGSFHCVCBEG 901
QY 353 FKLLTDEKSCQDVDECSLDRTCD-HS-CINHPGTACACNRYTL-YGTHCGDNTNECSI 409
Db 902 FSISADGRTCEDIDECVNTVCDSHGFCNDTAGSFRLCVQGFQAPDQGGVQVDMNECEL 961
QY 410 NNGGC-OCVCVNTVGSYECOC---HPCY-----KLNHNKKDC---Y 443
Db 962 LSGVCGEAFCEVGEFGLVCADENQESYPMTCQCRSVTEGSDVDRQPREKKECYKL 1021
QY 444 EVKGLLPTSVSPVSLH--CGKSGGGDGCFLRCHSGIHLSSDVTIRTSTVKLNKRGKCS 501
Db 1022 NDASLDNVLAENVTKQECCTSGAGW----- 1049
QY 502 LKNAELPP-----EGLRPALPKHSVSVKESFR-----YNLTCSSCK 538
Db 1050 -DNCEIFPCPVQGTAEFTMCPRGKGLVPAGESSYDTGGENYKDADECLLFGEEICKNGY 1108
QY 539 QVPGAPCRPSTPKEMFTVFELETNQKEVTAACDLSC-VKRTKRLRAIKR-LRKAVER 598
Db 1109 CLINTQPG-----YECYCKQGTYYDPVKLQC----- 1133
QY 599 EQFHQLQSGNMLDVAKKPPRTSERAESCGVGOGGHAENOCVS-----CRAGTYIDGA 650
Db 1134 -----FDMDECOQP-----NSC-----IDGQCWNTEGSYNCFTCPWVLDAS 1170
QY 651 RERCILCPNPTFQNESQMTCEBC-----FRPNSSALYK-----PEANWMBE 693
Db 1171 EXRCVQPTSENEQTEEDVDVYQDLQWEHLSEYVCSRLVKGKQTTYECCLLYGEAWGM-Q 1229
QY 694 CGGLCQPGSEYSDGFAFCQLCALGTGFOPEAGRTSCF-----PCGGGLATKHQ 740
Db 946 -----PCGGGLATKHQ 740
```

```
Db 1230 C-ALCPMKD--SDYA--QLCNI---PVTGRRRPYGRALVDFBSQYGFETDVPVFCQR 1280
QY 741 GATSFQDCETRVQCSPOHFFYNTTHRCIRCPVG-----TYQFEGKNCVSCPNTTIT 793
Db 1281 FLNMFEEQLAE-EC--GILNGCNGRCVVRVHEGYTCCDFDGYHLDMAKRTCDVV--NECS 1335
QY 794 DFGSTNITQCNRRCGGELGDF-----TGYES--PNY--PGNYPANTE 834
Db 1336 ELNNRMSL--CKNAKCNITGSKYKCLCLPGYIPSDKFNKYCTPLNSALNLD 1383

RESULT 12
LTBL_MOUSE
ID LTBL_MOUSE STANDARD: PRT: 1713 AA.
AC O8CGL9; O88349; Q8BNW7; O8C7P5; Q8CIR0;
DT 15-SEP-2003 (Rel. 42, Created:
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Latent transforming growth factor beta binding protein, isoform 1L
DE precursor (LTBP-1) (Transforming growth factor beta-1 binding protein
DE 1) (TGF-beta1-BP-1).
CN LTBP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus
OX NCBI_TaxID=10090;
RN 1;
RP SEQUENCE FROM N.A.
RX PubMed=12711388;
RA Noguera I., Obata H., Gualandris A., Cowie P., Rifkin D.B.;
RT "Molecular cloning of the mouse Ltbp-1 gene reveals a tissue specific
RT expression of alternatively spliced forms.";
RL Gene 306.3:41 (2003);
RN 2;
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J;
RA Weiskirchen R., Moser M., Guenther K., Weiskirchen S., Gressner A.M.;
RT "The murine latent transforming growth factor-beta binding protein
RT (Ltbp-1) is alternatively spliced, and maps to a region syntenic to
RL human chromosome 2p21-22.";
RN 3;
RP SEQUENCE OF 1112-1713 FROM N.A.
RX STRAIN=C57BL/6J;
RA Tissue=Aorta, Liver, and Vein;
RN 4;
RP PubMed=12466951;
RA Okazaki Y., Furuno Y., Kasukawa T., Adachi J., Bono H., Kondo S.;
RA Nishida I., Otsu N., Sato K., Suzuki H., Yamakawa I., Kiyosawa H.;
RA Yagi K., Toraru Y., Hasegawa Y., Nogami A., Schorbach C., Gotohori T.,
RA Badarrelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.;
RA Schriell L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.;
RA Blake J.A., Bradt D., Brusci V., Chothia C., Corbani J.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.;
RA Grimmond S., Gustincich S., Hirokawa N., Jackson J.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedziarski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.;
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petravsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.;
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Sempie C.A., Setou M., Shimada K.,
RA Sultana R., Takeraka Y., Taylor M.S., Teasdale R.D., Tomita K.,
RA Wilmberg J.G., Wyrshaw-Boris A., Wang Y., Watanabe Y., Wells C.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayashi N.,
RA Hirozane-Kikukawa I., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Iotani K., Ishii Y., Itoh M., Kagawa T.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
```



```

Db      908 DIDECAQVRLHC-SQRCENTEGSFLCVCPAGFMASEEETGNCIDVDDECLRPMDCMDGRCI 966
QY      100 H-----DCINPGNYRCT-CFDFMLAH 121
Db      967 NTAGAFCEYCDSGYRMSRGYCEDIDELKFTCTPEEQCVNTPSQYCVPTCEGFR--- 1023
QY      122 DGHN--CLVDDECLNNGGQHTCVNMVMSYSCCKKEGFFLSDNOHTC--IHRSEGLSC 177
Db      1024 -GMNGQCLVDDECLKPVCTNGSCTNLEGSYCMCHRGYSPTPDHRHCQDIDECQCNLC 1082
QY      178 MNKHGCHSHICKEAPRGVACRPGFEL--AKNORDCILTCNVHNGGQCH-SCDSTADG 234
Db      1083 MNGQ-----CRNTD-GSFRCTGGQYQLSAKQCEDIDECHE-HHLCSHGQCRNTSGS 1134
QY      235 PEGSCHPYQMYMTDGRSCLEREDTVLEVTESNTTSVVDGKVRKRRLJMETCAVNNGCCD 294
Db      1135 FQVCVQGYRASVLGDHC---ED---INE-----CLEDSVCQGG-- 1168
QY      295 RTCKDTSTGVHSCSPVGFLLDCKTKDIDECQTRNGGCDHF--CKNIVUSFDCGCKKG 352
Db      1169 -DCINTAGSYDCTCPDGFQNL-DNKCQDINBC-AQPLGCGHGECLATQGSFHCVCQG 1225
QY      353 FKLLTDEKSCQVDECSLDRCTD-HS-C-NHRTGTACACNRGYTL-VGPTHGCDTNECS 409
Db      1226 FS-SAGRTCEDIDECVNTVCDSHGFCNTAGSPRLCYQGFQAPQDQGCYVNVGEL 1285
QY      410 NNGGC-QQVCVNTVGYEQC---HPGY-----KLHNKKQC---V 443
Db      1286 LSGVCGEAFCEVNEGFLCVACDENCEYSPMTGQCSRVTEDSGVDROPRECKEYCNL 1345
QY      444 EVKGLLPTSVSRVSLH--CGKSGGDDGCLFCHSGIHLSSDVTIRTSTVFKLNEGKCS 501
Db      1346 NDASJCDNVLPVNTVQECCTSGAGWG----- 1373
QY      502 LKNAELFP-----EGLRALPKHSHSVKESFR-----YVNLTCSSGK 538
Db      1374 -DNCEIFPCPVQGTAEFTMCPRGKGLVPAGESSYDTGSENYKQADCEILLGEEICKNGY 1432
QY      539 QVPSAGRSTPKEMETITVEFLETNQKVTASCDLSCLVKTETKRLKRLTKLKAHVR 598
Db      1433 CLNTQSG-----YECYKQGYYYPVVLQC----- 1457
QY      599 EQPHQLSGMNLVDVAKPPRTSERQAESCQVGGQHAENQVS-----CRAGTYDGA 650
Db      1458 -----FQWDECCDP-----NSC-----IDGQCVTEGSSYNCFCFTHPWLDA 1494
QY      651 RERCILCPNGTFQNEGQWCTCP-----PRPNSGALK-----PERANNSE 693
Db      1495 ERKCVQFTESNEQIEETDYVQLCNEHLSEYVCSRPLVGRKQTYTECCCLYGEAWGM-Q 1553
QY      694 CGGLCPGYSADGAPCALGTPQBPAGTSCF-----PCGGGLATKHQ 740
Db      1554 A-ALCPMKD--SDYA--QLCNI-----PVTGRRPYGRDALVDFSEQVGPETDPYFQDR 1604
QY      741 GATSPGDCETRVQCSQGHFYNTTHRCIRCPVG-----TYQEPFGKNNCVSCFNGTTT 793
Db      1605 FLMSFEELQAE-EC--GILGNGCNGRCVRVHEGYTCDFGQVHLDMAKMTCDV--NECS 1659
QY      794 DFDGNTITQCKNRRCGGELGDF-----TGYES--PNY--PCNYPANTE 834
Db      1660 ELNRMSL--CRNAKCNTEGYSKCLCPGLFPSPDKPNYCTPLNSALNLD 1707

RESULT 13
LTBI RAT
ID LTBI RAT STANDARD; PRAT: 1712 AA.
AC Q00918;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Latent transforming growth factor beta binding protein 2 precursor
DE (LTBP-1) (Transforming growth factor beta-1 binding protein: 1) (TGF-

```

```

DE      betai-BP-1) (Transforming growth factor beta-1 masking protein, large
DE      subunit).
GN      LTSP1.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX      NCBI_TaxID=10116;
RN      [.]
RS      SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX      MEDLINE=91062373; PubMed=2247454;
RA      Tsuji T., Okada F., Yamaguchi K., Nakamura T.;
RT      "Molecular cloning of the large subunit of transforming growth factor
RT      type beta masking protein and expression of the mRNA in various rat
RT      tissues."
RL      Proc. Natl. Acad. Sci. U.S.A. 87:8835-8839(1990).
CC      -1- SUBUNIT: THE LARGE LATENT COMPLEX OF TGF-BETAL FROM PLATELETS IS
CC      COMPOSED OF THE TGF-BETAL MOLECULE NONCOVALENTLY ASSOCIATED WITH
CC      A MASKING PROTEIN CONSISTING OF A DISULFIDE-BONDED COMPLEX OF A
CC      DIMER OF THE N-TERMINAL PROPEPTIDE OF THE TGF-BETAL PRECURSOR AND
CC      A THIRD COMPONENT DENOTED TGF-BETAL-BP (OR MP LARGE SUBUNIT).
CC      TGF-BETAL-BP DOES NOT BIND DIRECTLY TO ACTIVE TGF-BETAL.
CC      -1- SUBCELLULAR LOCATION: Secreted.
CC      -1- SIMILARITY: Contains 18 EGF-like domains.
CC      This Swiss-Prot entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC      The European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
DR      EMBL; M5543; AAA2235.1; --
DR      PIR; A38261; A38261.
DR      HSP; P.6139; 1FSB.
DR      InterPro; IPR003152; Asx hydroxyl.
DR      InterPro; IPR001881; EGF_Ca.
DR      InterPro; IPR006259; EGF-like.
DR      InterPro; IPR002212; FGF-like-associ.
DR      Pfam; PF00005; EGF; 16.
DR      Pfam; PF00693; TB; 4.
DR      SMART; SM00179; EGF_CA; 13.
DR      ProSITE; PS00010; ASX HYDROXYL; 13.
DR      ProSITE; PS00022; EGF_1; 2.
DR      ProSITE; PS01186; EGF_2; 10.
DR      ProSITE; PS01187; EGF_CA; 15.
KW      Growth factor binding; Repeat; EGF-like domain; Hydroxylation; Signal.
FT      SIGNAL 1 20 POTENTIAL.
FT      PROPEP 21 736 POTENTIAL.
FT      CHAIN 737 1577 LATENT TRANSFORMING GROWTH FACTOR BETA
FT      B-INDING PROTEIN 1.
FT      SITE 734 736 CLEAVAGE (POTENTIAL).
FT      SITE 1575 1577 CLEAVAGE (POTENTIAL).
FT      PROPEP 1578 1712 POTENTIAL.
FT      DOMAIN 181 213 EGF-LIKE 1.
FT      DOMAIN 391 423 EGF-LIKE 2.
FT      REPEAT 551 604 INTERNAL REPEAT 1.
FT      DOMAIN 618 658 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT      REPEAT 671 721 INTERNAL REPEAT 2.
FT      DOMAIN 865 966 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT      DOMAIN 907 949 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT      DOMAIN 949 989 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT      DOMAIN 990 1029 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT      DOMAIN 1030 1070 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT      DOMAIN 1071 1111 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT      DOMAIN 1112 1152 EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).
FT      DOMAIN 1153 1193 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
FT      DOMAIN 1194 1235 EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
FT      DOMAIN 1236 1277 EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).
FT      DOMAIN 1278 1320 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
FT      REPEAT 1340 1392 INTERNAL REPEAT 3.
FT      DOMAIN 1415 1457 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
FT      DOMAIN 1458 1498 EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).

```

FT	REPEAT	1517	1568	INTERNAL REPEAT 4.
FT	DOMAIN	1612	1652	EGF-LIKE 17.
FT	DOMAIN	1653	1697	EGF-LIKE 18. CALCIUM-BINDING (POTENTIAL).
FT	D:SULFID	185	195	BY SIMILARITY.
FT	D:SULFID	189	201	BY SIMILARITY.
FT	D:SULFID	203	212	BY SIMILARITY.
FT	D:SULFID	395	405	BY SIMILARITY.
FT	D:SULFID	399	411	BY SIMILARITY.
FT	D:SULFID	413	422	BY SIMILARITY.
FT	D:SULFID	622	633	BY SIMILARITY.
FT	D:SULFID	628	642	BY SIMILARITY.
FT	D:SULFID	644	657	BY SIMILARITY.
FT	D:SULFID	869	881	BY SIMILARITY.
FT	D:SULFID	876	890	BY SIMILARITY.
FT	D:SULFID	892	905	BY SIMILARITY.
FT	D:SULFID	911	923	BY SIMILARITY.
FT	D:SULFID	918	932	BY SIMILARITY.
FT	D:SULFID	934	947	BY SIMILARITY.
FT	D:SULFID	953	964	BY SIMILARITY.
FT	D:SULFID	959	973	BY SIMILARITY.
FT	D:SULFID	976	988	BY SIMILARITY.
FT	D:SULFID	994	1005	BY SIMILARITY.
FT	D:SULFID	1000	1014	BY SIMILARITY.
FT	D:SULFID	1017	1028	BY SIMILARITY.
FT	D:SULFID	1034	1045	BY SIMILARITY.
FT	D:SULFID	1040	1054	BY SIMILARITY.
FT	D:SULFID	1056	1069	BY SIMILARITY.
FT	D:SULFID	1075	1086	BY SIMILARITY.
FT	D:SULFID	1081	1095	BY SIMILARITY.
FT	D:SULFID	1097	1110	BY SIMILARITY.
FT	D:SULFID	1116	1127	BY SIMILARITY.
FT	D:SULFID	1122	1136	BY SIMILARITY.
FT	D:SULFID	1138	1151	BY SIMILARITY.
FT	D:SULFID	1157	1169	BY SIMILARITY.
FT	D:SULFID	1164	1178	BY SIMILARITY.
FT	D:SULFID	1180	1192	BY SIMILARITY.
FT	D:SULFID	1198	1210	BY SIMILARITY.
FT	D:SULFID	1204	1219	BY SIMILARITY.
FT	D:SULFID	1221	1234	BY SIMILARITY.
FT	D:SULFID	1240	1252	BY SIMILARITY.
FT	D:SULFID	1246	1261	BY SIMILARITY.
FT	D:SULFID	1263	1276	BY SIMILARITY.
FT	D:SULFID	1282	1294	BY SIMILARITY.
FT	D:SULFID	1289	1303	BY SIMILARITY.
FT	D:SULFID	1305	1319	BY SIMILARITY.
FT	D:SULFID	1419	1432	BY SIMILARITY.
FT	D:SULFID	1427	1441	BY SIMILARITY.
FT	D:SULFID	1443	1456	BY SIMILARITY.
FT	D:SULFID	1462	1473	BY SIMILARITY.
FT	D:SULFID	1468	1482	BY SIMILARITY.
FT	D:SULFID	1484	1497	BY SIMILARITY.
FT	D:SULFID	1616	1627	BY SIMILARITY.
FT	D:SULFID	1622	1636	BY SIMILARITY.
FT	D:SULFID	1638	1651	BY SIMILARITY.
FT	D:SULFID	1657	1672	BY SIMILARITY.
FT	D:SULFID	1667	1681	BY SIMILARITY.
FT	D:SULFID	1693	1696	BY SIMILARITY.
FT	CARBOHYD	339	370	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	370	416	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	416	416	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	612	612	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1042	1042	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1242	1242	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1357	1357	N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE	1712 AA;	186598 MW; 650BCEAA691FD134 CRC64;	

Query Match 10.8%; Score 592; DB 1; Length 1712;  
Best Local Similarity 24.3%; Pred. No. 1,le-31;  
Matches 230; Conservative 102; Mismatches 305; Indels 310; Gaps 52;

QY	45	DVDECAQGLDDHADALCONTPTSYKCSCKPGY--QGEGRCQCEIDDEC-----GNEL 94
DB	907	DIDECAQAOHLCC-SQRCENTEGSFUICIPAGFIASEGNSCIDVDECLRPVDCRGRCI 965

QY	95	N-----GGC-----VHDCLNIPGNYRCT-CFCDGFMJAH 121
DB	966	NTAGAFRCVCDGSRMSRRGHGCD-DECLTPGCPREEQVNSPGSYQCVPTCEGR--- 1022
QY	122	DGHN--CLDYDECLNNGGCOHTCVNMVMSYECCKEKFELSDNQHTC--IHRSEGLSC 177
DB	1023	-GNGQCJVDDECLQPKVCTNGSCTNLEGSYKCSCHKYSPTPCHRHCDQIDECQGNJC 1081
QY	178	VXKDHGCHICKAPRGSVACERPGFEL--AKNQRDCILTCNH3NGGCGH--SCDIDAOG 234
DB	1082	MNGO-----CKNTD--GSPRCTGGQYQLSAKDQCEDIDCEHRRH-LSHGQCRTEBS 1133
QY	235	PEGSHCPQVQCHTQGRSCLEREDTVLEVTESENTTSVVGDGRVKRRLJLMMETCAVNGGCD 294
DB	1134	FQCLNQGYEASVLGDHC--ED-----INE-----CLEDSSVCQGG-- 1167
QY	295	RTCKDTSTGVHCSCPVGFLLQDGKTKCIDBCOTRNGGCDHF--CKNIUGSFFDCCKKG 352
DB	1168	-DCINTAGSYDCTCPDGLQIN-DNKQCODINEC-AQPLGCAHGECLNTQGSFHCYCEG 1224
QY	353	FKULTDEKSCQVDECSJDRTCO-HS-CINHPCTFACACNRGYTL-VGFTHGCTTVECSI 409
DB	1225	FSISAGRTCEDIDECUNNTVCDSHGFCNTAGSFCLCYGFCAPQDGGCVQVNECEL 1284
QY	410	NNGGC-QQVCNTVGSYECOCFPGYKLH-----WNKDC---V 443
DB	1285	LSGVGGEAFCEVNEGSGFLVCADENQESPMTCQSRATEDSGVDRQPKKEKCEYVNL 1344
QY	444	EVKGLFTSVSPRVSJH--CGKSGGDDGCFLRCHSGIHLSOVTTTTSVTFKLNEGKS 591
DB	1345	NDASLDCENVLAPNVTKQBCCTSGAGW-----YVNLTCSSGK 538
QY	502	LKNAELFP-----EGLRALPEKHSVKESFR-----YVNLTCSSGK 538
DB	1373	-DNCEIFPCVPQGTABEFSEBMCPRGKGFVPAGESSEYETGGENYKDADECLLPGEI 1431
QY	539	QVFGAPGRPTPKEMFIVTEFELETNOKEVTASCDLSICIVKTEKRLKXALRLKXAVHR 598
DB	1432	C.LNTQGYECYCKE-----GTYPDPVKLQC----- 1456
QY	599	EQFHLQLSGNLDVAKKPTSERQAESGVGGHAENQCVS-----CRAGTYDDGA 650
DB	1457	-----FMDDECQDP-----NSC-----IDGQCVNTEGYNCFCTHPVVDAS 1493
QY	651	RERCILCPNGTFOKEGQMTCEPC-----PRGNSGALKT-----PEANVMSR 693
DB	1494	EKRCVQPTESNBOIEETDVVDLQWEHLSEYVCSRPJGKQCTTYTECCLYGEANGM-Q 1552
QY	694	CGGLQCFGEYSADGAPQCLCALGTTPQEAQRTSCF-----PCGGGLATKHQ 740
DB	1553	C-ALCEMKD--SDQYA--QLCNI-----PVTGRRRPYGRDALVDFSEQYGPETDPYFIQDR 1603
QY	741	GATSFQDCETRVQCSSECHFYNTTHRCICRPVG-----TYQPEGKNKNCUSCPQNTIT 793
DB	1604	FLANSFEELQAE-EC--GILNGCENGRCVRQESYTCDFDGHLDMAKTCVDV--NECS 1658
QY	794	DFDGSNITCKNRRCGGELGDF-----GYIES--PNVPGNYNPANT 832
DB	1659	ELANRMSL--CKNAKCINTEGSKYKCVCLFGYVPSDKPNY--CTPLNT 1701

RESULT 14  
FBL2 MOUSE  
ID FBL2 MOUSE STANDARD; PRT: 1221 AA.  
AC P37855; Q3WU12;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE FBLin-2 precursor.  
CN FBLN2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-35.  
 RC TISSUE=Fibroblast;  
 RX MEDLINE=94064787; PubMed=82451130;  
 RA Pan T.-C., Sasaki T., Zhang R.-Z., Faessler R., Timpl R., Chu M.-J.;  
 "Structure and expression of fibulin-2, a novel extracellular matrix  
 protein with multiple EGF-like repeats and consensus motifs for  
 calcium binding.";  
 RT J. Cell Biol. 123:1269-1277(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RX MEDLINE=99337686; PubMed=10406956;  
 RA Graessel S., Sicot F.-X., Gotta S., Chu M.-L.;  
 "Mouse fibulin-2 gene. Complete exon-intron organization and promoter  
 characterization.";  
 RT Eur. J. Biochem. 263:471-477(1999).  
 RN [3]  
 RP DEVELOPMENTAL STAGE.  
 RX PubMed=850569;  
 RA Zhang H.-Y., Timpl R., Sasaki T., Chu M.-L., Ekblom P.;  
 "Fibulin-1 and fibulin-2 expression during organogenesis in the  
 developing mouse embryo.";  
 RT Dev. Dyn. 205:348-364(1996).  
 RN [4]  
 RP BINDING TO LAMA2.  
 RX PubMed=1002829;  
 RA Talts J.F., Andac Z., Goshring W., Brancaccio A., Timpl R.;  
 "Binding of the G domains of laminin alpha1 and alpha2 chains and  
 perlecan to heparin, sulfatides, alpha-dystroglycan and several  
 extracellular matrix proteins.";  
 RT EMBO J. 18:863-870(1999).  
 RN [5]  
 RP DOWN-REGULATION BY GLUCOCORTICOIDS.  
 RX PubMed=11737251;  
 RA Gu Y.-C., Talts J.F., Gullberg D., Timpl R., Ekblom M.;  
 "Glucocorticoids down-regulate the extracellular matrix proteins  
 fibronectin, fibulin-1 and fibulin-2 in bone marrow stroma.";  
 RT Eur. J. Haematol. 67:176-184(2001).  
 CC -!- FUNCTION: ITS BINDING TO FIBRONECTIN AND SOME OTHER LIGANDS IS  
 CC CALCIUM DEPENDENT.  
 CC -!- SUBUNIT: Homotrimer; disulfide-linked. Interacts with LAMA2.  
 CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Comment=Additional isoforms seem to exist;  
 CC Name=1;  
 CC IsoId=P37889-1; Sequence=Displayed;  
 CC Name=2; Synonyms=EGF3-less;  
 CC IsoId=P37889-2; Sequence=VSP 001391;  
 CC -!- TISSUE SPECIFICITY: Component of both basement membranes and other  
 CC connective tissues.  
 CC -!- DEVELOPMENTAL STAGE: The differential expression of the fibulin  
 CC family contributes to the formation of molecularly distinct  
 CC extracellular matrices already during early developmental stages  
 CC of a large number of tissues.  
 CC -!- INDUCTION: Glucocorticoids suppressed mRNA expression and protein  
 CC synthesis.  
 CC -!- SIMILARITY: Belongs to the fibulin family.  
 CC -!- SIMILARITY: Contains 3 anaphylatoxin-like domains.  
 CC -!- SIMILARITY: Contains 11 EGF-like domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X75285; CAAS3040.1;  
 DR EMBL; AF135253; AAD34456.1; -;  
 DR EMBL; AF135253; AAD34456.1; -;

DR EMBL; AF135239; AAD34456.1; JOINED.  
 DR EMBL; AF135240; AAD34456.1; JOINED.  
 DR EMBL; AF135241; AAD34456.1; JOINED.  
 DR EMBL; AF135242; AAD34456.1; JOINED.  
 DR EMBL; AF135243; AAD34456.1; JOINED.  
 DR EMBL; AF135244; AAD34456.1; JOINED.  
 DR EMBL; AF135245; AAD34456.1; JOINED.  
 DR EMBL; AF135246; AAD34456.1; JOINED.  
 DR EMBL; AF135247; AAD34456.1; JOINED.  
 DR EMBL; AF135248; AAD34456.1; JOINED.  
 DR EMBL; AF135249; AAD34456.1; JOINED.  
 DR EMBL; AF135250; AAD34456.1; JOINED.  
 DR EMBL; AF135251; AAD34456.1; JOINED.  
 DR EMBL; AF135252; AAD34456.1; JOINED.  
 DR FIR; A49457; A49457.  
 DR HSP; P0736; IAPQ.  
 DR MGI; 95488; Pbln2.  
 DR InterPro: IPR000020; Anaphylatoxin.  
 DR InterPro: IPR000152; Asx\_hydroxyl.  
 DR InterPro: IPR001861; EGF\_Ca.  
 DR InterPro: IPR006209; EGF\_like.  
 DR Pfam: PF01821; ANATO; 2.  
 DR Pfam: PF00308; EGF; 6.  
 DR SMART: SM01104; ANATO; 3.  
 DR SMART: SM00179; EGF\_Ca; 9.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 5.  
 DR PROSITE: PS01177; ANAPHYLATOXIN\_1; 3.  
 DR PROSITE: PS01178; ANAPHYLATOXIN\_2; 3.  
 DR PROSITE: PS00022; EGF\_1; FALSE\_NEG.  
 DR PROSITE: PS01196; EGF\_2; 5.  
 DR PROSITE: PS01187; EGF\_Ca; 10.  
 KW Signal; Glycoprotein; Extracellular matrix; Plasma; EGF-like domain;  
 KW Calcium-binding; Alternative splicing; Repeat.  
 FT SIGNAL 1 26  
 FT CHAIN 27 4221 FIBULIN-2.  
 FT DOMAIN 27 434 N.  
 FT DOMAIN 27 434 SUBDOMAIN NA (CYS-RICH).  
 FT DOMAIN 177 434 SUBDOMAIN NB (CYS-FREE).  
 FT DOMAIN 435 477 ANAPHYLATOXIN-LIKE 1.  
 FT DOMAIN 478 510 ANAPHYLATOXIN-LIKE 2.  
 FT DOMAIN 511 543 ANAPHYLATOXIN-LIKE 3.  
 FT DOMAIN 594 635 EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 669 708 EGF-LIKE 2.  
 FT DOMAIN 709 755 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 756 800 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 801 846 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 847 894 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 895 937 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 938 979 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 1019 1061 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 1062 1106 EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 1111 1221 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).  
 FT SITE 421 423 DOMAIN III.  
 FT SITE 421 423 CELL ATTACHMENT SITE (POTENTIAL).  
 FT DISULFID 435 462 BY SIMILARITY.  
 FT DISULFID 436 469 BY SIMILARITY.  
 FT DISULFID 449 470 BY SIMILARITY.  
 FT DISULFID 479 508 BY SIMILARITY.  
 FT DISULFID 492 509 BY SIMILARITY.  
 FT DISULFID 511 535 BY SIMILARITY.  
 FT DISULFID 512 542 BY SIMILARITY.  
 FT DISULFID 525 543 BY SIMILARITY.  
 FT DISULFID 598 610 BY SIMILARITY.  
 FT DISULFID 606 619 BY SIMILARITY.  
 FT DISULFID 621 634 BY SIMILARITY.  
 FT DISULFID 673 683 BY SIMILARITY.  
 FT DISULFID 673 692 BY SIMILARITY.  
 FT DISULFID 694 707 BY SIMILARITY.  
 FT DISULFID 713 726 BY SIMILARITY.  
 FT DISULFID 723 735 BY SIMILARITY.  
 FT DISULFID 742 754 BY SIMILARITY.  
 FT DISULFID 805 818 BY SIMILARITY.  
 FT DISULFID 812 827 BY SIMILARITY.



```

FT DISULFID 833 845 BY SIMILARITY.
FT DISULFID 899 912 BY SIMILARITY.
FT DISULFID 936 921 BY SIMILARITY.
FT DISULFID 923 936 BY SIMILARITY.
FT DISULFID 942 954 BY SIMILARITY.
FT DISULFID 950 963 BY SIMILARITY.
FT DISULFID 965 978 BY SIMILARITY.
FT DISULFID 984 993 BY SIMILARITY.
FT DISULFID 989 1002 BY SIMILARITY.
FT DISULFID 1004 1017 BY SIMILARITY.
FT DISULFID 1023 1035 BY SIMILARITY.
FT DISULFID 1031 1044 BY SIMILARITY.
FT DISULFID 1046 1060 BY SIMILARITY.
FT DISULFID 1066 1079 BY SIMILARITY.
FT DISULFID 1073 1088 BY SIMILARITY.
FT DISULFID 1093 1105 BY SIMILARITY.
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 497 497 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 737 737 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1072 1072 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPIC 709 755 Missing (in isoform 2).
FT FTID-VSP 001391.
FT CONFLICT 140 159 HSRKKAAGTHVHSSCRAC -> TVAVSICHVYAPPLLP
FT GF (IN REF. 2).
FT CONFLICT 348 348 S -> L (IN REF. 2).
FT CONFLICT 507 507 Q -> QO (IN REF. 2).
FT CONFLICT 1102 1102 Q -> E (IN REF. 2).
SQ SEQUENCE 1221 AA; 131818 MW; 87DB2A10A8FDC45F CRC64;

Query Match
Best Local Similarity 29.0%; Pred. No. 7.1e-31;
Matches 157; Conservative 66; Mismatches 182; Indels 137; Gaps 29;

QY 44 EDVDECAQGLDCHADALCONTPYSKC-----SKPGY-QEGGROCEDIDCGNELNGG 97
DB 708 EDQDECLMGTHDCSKQFCVNTUGSYCVNHTVLCAGYILNAHRKCVDINECVLDLH-T 766

QY 98 C--VHDCMLTPGNRYC-----TCDFGMLAHGHNCCDVDECLNNGGCG--HTCVNVMGS 149
DB 767 CTRAHCNVNTPGSGFQYKALTCFPGVLT-DG-ECTDVECVT-GTHNCOAGSCQNTKGS 824

QY 150 YEC-----CKEGFPLSDNQHTCHRSR-----EGLSCMNKHGCHSHCKEAPRGSV 196
DB 825 FYCAQRQRCDMG-FLQDPEGNVCVDINECTSLLEPCRSFGSCINTVG--SYTCQNP---- 877

QY 197 ACSCRGFELAKNORDC--ILTCNHNNGGCOHS--CDDTAGPECCSCHPCYXQHTDGRSC 252
DB 878 -LYCGRGYHANESEGVNVECTGVHRCGEQLCYNLPGSVRCCKGFGORDAFGRTC 936

QY 253 LEREDTVLEVTESNTT-SWVDGDKRVKRLMETCAVYNGG-CORTCKDTSTGVHSCPYG 311
DB 937 ID-----VNECWSPGRLCQHTCENTPGSVRCSAAG 968

QY 312 FTQLDGKTKCIDECQTRNGGCHDFCKNLVSGFDCGKKKGFKLLTDEKSCQDVIECSLD 371
DB 969 FLAADGKHGCVNECTETPR--CSQECANIYGYQYCRQYGLAEDGHTCTCIDCECAQ3 1026

QY 372 R--TCDSHCNHHGTTFACAC-NRGYTLV-GFTHCGDINECSINNGGCG--QVCNNTVGSY 425
DB 1027 AGILCTFRCVNPGSVQACPEGYTNMANGRSCKDLDECALGTHNCSAECHNLCQSF 1086

QY 426 EC---QCHPOY-----KLHNMNKDCVEY-----KGLJL----- 449
DB 1087 RCLRFDCPPNVRVRSQTKCERTTQDITECQTSFARITHYQLNFQGLLVPAHIFRIGFA 1146

QY 450 PTVSVRSLVHCKSGGSGGDFCRHSGIHLSSDVTIRTSVTFKLN--EGKCSLKNABL 507
DB 1147 PAFAGDTISITRK-GNEEGYF-----VTRRLNAYTGVVSLQSRVSL 1186

QY 508 FP 509
DB 1187 EP 1188

```

## RESULT 15

```

FB11_HUMAN
ID FB11_HUMAN STANDARD; PRT; 703 AA.
AC P23142; P23143; P23144; P37888; Q8TBH6; Q8HB05; C9U3R4; C9UH41;
DT C1-NOV-1991 (Rel. 23, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Fibulin-1 precursor.
DE Fibulin-1 precursor.
GN FB1N1.
CS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A; B AND C).
RX MEDLINE=9100426; PubMed=2269669;
RA Argaves W.S., Tran H., Burgess W.H., Dickerson K.;
RT "Fibulin is an extracellular matrix and plasma glycoprotein with
RT repeated domain structure.";
RL J. Cell Biol. 111:3155-3164(1990).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM D), TISSUE SPECIFICITY, AND INTERACTION
RP WITH FN1 AND FGB.
RX PubMed=9106159;
RA Tran H., Mattei M., Godyna S., Argaves W.S.;
RT "Human fibulin-ID: molecular cloning, expression and similarity with
RT SL-5 protein, a new member of the fibulin gene family.";
RL Matrix Biol. 15:479-493(1997).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM D).
RX MEDLINE=99253993; PubMed=10318851;
RA Krichevsky A.M., Metzger E., Rosen H.;
RT "Translational control of specific genes during differentiation of
RT HL-60 cells.";
RL J. Biol. Chem. 274:14295-14305(1999).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM C).
RA Gu J.R., Wan D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,
RA Qin W.X., Huang Y., Qiu X.K., Qian L.F., He L.F., Li H.N., Yu Y.,
RA Yu J., Han L.H.;
RT "Novel Human cDNA clones with function of inhibiting cancer cell
RT growth.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=20057165; PubMed=10591208;
RA Dunham T., Hunt A.R., Collins C.E., Bruskewich R., Beare D.M.,
RA Clapp M., Smink J.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark S.,
RA Clegg S.N., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Corcoran D., Corby K.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA Dham P.D., Dockree K., Dodsworth S.J., Durbin R.M., Ellington A.G.,
RA Evans K.L., Fey J.M., Fleming K., French J., Garner A.A.,
RA Gilbert J.G.R., Goward M.E., Graham D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tamlyn G., Heathcott R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kershaw C., Kimberley A.M., King A.,
RA Jald G.K., Langford C.F., Leversha X.A., Lloyd C., Lloyd D.M.,
RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann C.T.,
RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillips B.J.C.,
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers J., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.,
RA Soderlund C., Spragon L., Steward C.A., Sturton J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis C.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming J.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki K., Mitsuyama S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Doran A., Fang F., Fu Y., Hu P., Hua A., Kenton S., La H., Lac H.I.,

```

RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Par H.,  
 RA Phan S., Qi S., Qian Y., Ray L., Ren C., Shaull S., Sloan D., Song L.,  
 RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,  
 RA Zhan M., Zhang G., Chissoe S., Murray J., Miller N., Minx P.,  
 RA Fulton R., Johnson D., Benis G., Bentley D., Bradshaw H., Bourne S.,  
 RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,  
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing J.,  
 RA Scheet P., Walker C., Wamsley A., Wohldmann P., Pepin K., Nelson J.,  
 RA Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,  
 RA Emanuel B.S., Shaikh T., Kurahashi H., Saitta S., Budarf M.L.,  
 RA McDermid H.B., Johnson A., Wong A.C.C., Morrow B.E., Edelmann L.,  
 RA Kim U.J., Shizuya H., Simon M.I., Dumanski J.P., Peyrard M., Keira C.,  
 RA Seroussi E., Fransoni I., Tapia I., Bruder C.E., O'Brien K.P.,  
 RA Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,  
 RA Tiliakou Y., Wright H.,  
 RT "The DNA sequence of human chromosome 22";  
 RL Nature 402:489-495(1999).  
 RN [6].  
 RP SEQUENCE FROM N.A. (ISOFORM C).  
 RC TISSUE=Brain;  
 RX MEDLINE=22398257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Altschul S.F., Zeng B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.J., Scheetz T.E.,  
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Frange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullatly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton S., Ketterman M., Yadan A., Rodrigues S., Sanchez A.,  
 RA Whiting X., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Jackson X.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield J.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [7].  
 RP SEQUENCE OF 1-26 FROM N.A.  
 RX PubMed=11829738;  
 RA Castoldi M., Chu M.-L.;  
 RT "Structural and functional characterization of the human and mouse  
 RT fibulin-1 gene promoters: role of Sp1 and Sp3";  
 RL Biochem. J. 362:41-50(2002).  
 RN [8].  
 RP SEQUENCE OF 30-44.  
 RX MEDLINE=89354537; PubMed=2527614;  
 RA Argaves W.S., Dickerson K., Burgess W.H., Ruoslahti E.;  
 RT "Fibulin, a novel protein that interacts with the fibronectin  
 RT receptor beta subunit cytoplasmic domain";  
 RL Cell 58:623-629(1989).  
 RN [9].  
 RP SELF-ASSOCIATION AND INTERACTION WITH FN1.  
 RX PubMed=1400330;  
 RA Balbona K., Tran H., Godyna S., Ingham K.C., Strickland D.K.,  
 RA Argaves W.S.;  
 RT "Fibulin binds to itself and to the carboxyl-terminal heparin-binding  
 RT region of fibronectin";  
 RL J. Biol. Chem. 267:20120-20125(1992).  
 RN [10].  
 RP POSSIBLE FUNCTION.  
 RX PubMed=7534784;  
 RA Roark E.F., Keene D.R., Haudenschild C.C., Godyna S., Little C.D.,  
 RA Argaves W.S.;  
 RT "The association of human fibulin-1 with elastic fibers: an  
 RT immunohistological, ultrastructural, and RNA study";  
 RL J. Histochem. Cytochem. 43:401-411(1995).  
 RN [11].  
 RP INTERACTION WITH FGB.  
 RX PubMed=7642629;  
 RA Tran H., Tanaka A., Litvinovich S.V., Meved L.V., Haudenschild C.C.,  
 RA Argaves W.S.;  
 RT "The interaction of fibulin-1 with fibrinogen. A potential role in  
 RT hemostasis and thrombosis";  
 RL J. Biol. Chem. 270:19458-19464(1995).  
 RN [12].  
 RP DEVELOPMENTAL STAGE.  
 RX PubMed=8737292;  
 RA Miosge N., Gotsch W., Sasaki T., Chu M.-L., Timpl R., Herken R.;  
 RT "The extracellular matrix proteins fibulin-1 and fibulin-2 in the  
 RT early human embryo";  
 RL Histochem. J. 26:109-116(1996).  
 RN [13].  
 RP INDUCTION.  
 RX MEDLINE=96133928; PubMed=8552629;  
 RA Clinton G.V., Rougeot C., Derancourt J., Roger P., DeRenne A.,  
 RA Godyna S., Argaves W.S., Rochefort H.;  
 RT "Estrogens increase the expression of fibulin-1, an extracellular  
 RT matrix protein secreted by human ovarian cancer cells";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:318-320(1996).  
 RN [14].  
 RP CALCITRIOL, SELF-ASSOCIATION, AND FN1-BINDING SITES.  
 RX PubMed=9278415;  
 RA Tran H., VanDusen W.J., Argaves W.S.;  
 RT "The self-association and fibronectin-binding sites of fibulin-1 map  
 RT to calcium-binding epidermal growth factor-like domains";  
 RL J. Biol. Chem. 272:22600-22606(1997).  
 RN [15].  
 RP ROLE IN TUMOR FORMATION AND INVASION.  
 RX PubMed=9393974;  
 RA Qing J., Maher V.X., Tran H., Argaves W.S., Dunstan R.W.,  
 RA McCormick J.J.;  
 RT "Suppression of anchorage-independent growth and matrix invasion and  
 RT delayed tumor formation by elevated expression of fibulin-1D in human  
 RT fibrosarcoma-derived cell lines";  
 RL Oncogene 15:2159-2168(1997).  
 RN [16].  
 RP INDUCTION.  
 RX PubMed=9811350;  
 RA Roger P., Pujol P., Lucas A., Baidet P., Rochefort H.;  
 RT "Increased immunostaining of fibulin-1, an estrogen-regulated protein  
 RT in the stroma of human ovarian epithelial tumors";  
 RL Am. J. Pathol. 153:1579-1588(1998).  
 RN [17].  
 RP ROLE IN TUMOR FORMATION AND INVASION.  
 RX PubMed=9466571;  
 RA Hayashita Y., Lucas A., Rougeot C., Godyna S., Argaves W.S.,  
 RA Rochefort H.;  
 RT "Estradiol and fibulin-1 inhibit motility of human ovarian- and  
 RT breast-cancer cells induced by fibronectin";  
 RL Int. J. Cancer 75:654-658(1998).  
 RN [18].  
 RP INTERACTION WITH NOV.  
 RX PubMed=9927866;  
 RA Perbal B., Martinier C., Sainson R., Werrin M., He B., Poizman B.;  
 RT "The C-terminal domain of the regulatory protein NOV is sufficient to  
 RT Query Match 10.1%; Score 567; DB 1; Length 703;  
 Best Local Similarity 28.8%; Pred. No. 1.9e-30;  
 Matches 159; Conservative 75; Mismatches 167; Indels 152; Gaps 29;  
 Qy 44 EDVDECAQGLDDCHADALCONTPTSYK----SCKPGYQ-GEGRCEIDECNENGGCC 98  
 Db 2:5 EDVDECC-TGSHSCRUGESCIN-VGSRFCORDSSCGTVELTDSNCKDIDECESIH-NC 273  
 Qy 99 VHP--CNPNGYRC-----TCDFGNLHAGHNCILVDCELENNGGCC-HTCVNMWSY 150  
 Db 274 LPPIFICNTLGSFRRCRKLQCKSGFIQALG-NCIDINECLISAPCIGTCTCNGESY 332  
 Qy 151 EG-----CCKEGFFLSDNQTCHPSEGLSCMNKDH-----GCSEHCKEAFRGQVA 197  
 Db 333 TCQKNVPCGRGY-----HUNEETRCVQVDECAPPAEPCGKGRVNSP-GSFR 381



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2003, 11:16:54 ; Search time 88 Seconds  
(without alignments)  
1801.907 Million cell updates/sec

Title: US-09-747-371-2

Perfect score: 5605  
Sequence: 1 MGVAGNRPGAAWVLLLL.....PRSFRLLRKSVSRFLRPYK 999

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5605	100.0	999	23 ABP69329	Human polypeptide
2	5605	100.0	999	23 AB705567	Breast cancer-asso
3	5605	100.0	999	23 AB880926	Human breast cancer
4	5605	100.0	999	23 ABG61893	Prostate cancer-asso
5	5605	100.0	999	24 ABR47407	Breast cancer-asso
6	5605	100.0	999	24 ABJ19815	Androgen-independe
7	5581	99.6	997	24 ABJ37049	Human breast cancer
8	5428	96.8	964	24 ABP56757	Human CEGP1 protei
9	5079	90.6	919	24 AAO16642	Human extracellular

# ALIGNMENTS

RESULT 1  
ABP69329  
IC ABP69329 standard; Protein; 999 AA.

AC ABP69329;

XX 20-JAN-2003 (first entry);

XX Human polypeptide SEQ ID NO 1376.

DE Human; genome mapping; gene therapy; food supplement; virus; fungus;  
cell-proliferative disorder; neurodegenerative disease; bacterial;  
Parkinson's disease; Alzheimer's disease; autoimmune disease;  
multiple sclerosis; diabetes; genetic disorder; wound; infection;  
arthritis; cytostatic; immunomodulator; noctropic; neuroprotective;  
anti-parkinsonian; antidiabetic; immunosuppressive; dermatological;  
haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;  
antiarthritic.

OS Homo sapiens.

XX WO200270539-A2.

XX 12-SEP-2002.

XX 05-MAR-2002; 2002WO-USC0595.

XX 05-MAR-2001; 2001US-C799451.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;

Mouse BCC2 ortholo  
Mouse CEGP1 protei  
Human epidermal gr  
Protein of NOVX 8b  
Human VEGF/FIBRILL  
Novel human epider  
Human LP283 protei  
Human POLY7 protei  
Human LP283 mature  
Amino acid sequenc  
Human LP283 splice  
Human POLY8 protei  
Human extracellular  
Novel human epider  
Amino acid sequenc  
Amino acid sequenc  
Protein of NOVX 8a  
Human EGF-Related  
Mouse EGF-like 1 p  
Human LP283 splice  
Human POLY5 protei  
Novel human epider  
Novel human diagno  
Amino acid sequenc  
Primate LP283. Ya  
Human breast-speci  
Human POLY5 protei  
Novel human epider  
Human MDDT polypep  
Breast cancer prot  
Human ovarian anti  
Amino acid sequenc  
Human 56739 amino  
Human LP283 splice  
Amino acid sequenc  
Human CDNA SEQ ID

PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;  
 PI Wehrman T, Wang J, Wang D, Drmanac RT;  
 XX WPI; 2002-759812/82.  
 DR N-PSDB; AB211546.  
 XX  
 XX New polynucleotides comprising sequences assembled from expressed  
 PT sequence tags (ESTs), useful for treating cell-proliferative, or  
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or  
 PT platelet or coagulation disorders -  
 XX  
 XX Claim 9; SEQ ID NO 1376; 1012pp + Sequence Listing; English.  
 XX  
 XX The invention relates to an isolated polynucleotide (\*) comprising a  
 CC nucleotide sequence selected from any of 948 sequences  
 CC (AB2111.9-AB212066) or their mature protein coding portion, active domain  
 CC coding protein or complementary sequences. The polynucleotides are useful  
 CC for identifying expressed genes or for physical mapping of human genome.  
 CC The encoded polypeptides (ABP6902-ABP6949) are useful as molecular  
 CC weight markers, as a food supplement, for generating antibodies, in  
 CC medical imaging, screening and diagnostic assays and for treating  
 CC cell-proliferative disorders (cancer), neurodegenerative diseases  
 CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple  
 CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid  
 CC disorders, platelet or coagulation disorders, wound, burns, incision,  
 CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,  
 CC parasitic), arthritis, etc.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 999 AA;

Query Match 100.0%; Score 5605; DB 23; Length 999;  
 Best Local Similarity 100.0%; Fred. No. 0;  
 Matches 999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGAGNRGGAAMVALLLLPPLLLLAGAVPPGRAGAGPOEDVDECAOGLDCHADA 60  
 DB 1 MGAGNRGGAAMVALLLLPPLLLLAGAVPPGRAGAGPOEDVDECAOGLDCHADA 60

QY 61 LCQNTPTSKYCKCPGYQEGRQCEDIDECGNELNGGCVHDCNLIPGNVRYCTCFDGMFLA 120  
 DB 61 LCQNTPTSKYCKCPGYQEGRQCEDIDECGNELNGGCVHDCNLIPGNVRYCTCFDGMFLA 120

QY 121 HDGNCLDVECLNNGGQHTCVNMGSYECCKEGFELSDNQTCTHRSEEGSCWNK 180  
 DB 121 HDGNCLDVECLNNGGQHTCVNMGSYECCKEGFELSDNQTCTHRSEEGSCWNK 180

QY 181 DRGSHICKEAPRGSVACERPGFELAKNQDRICILTCNHNNGGCGQSCDDTADGPECSC 240  
 DB 181 DRGSHICKEAPRGSVACERPGFELAKNQDRICILTCNHNNGGCGQSCDDTADGPECSC 240

QY 241 PQYKHTDGRSLEREDTVLETSNTTSVWDGKRVKRLMETCAVNNGCSDRTCKDT 300  
 DB 241 PQYKHTDGRSLEREDTVLETSNTTSVWDGKRVKRLMETCAVNNGCSDRTCKDT 300

QY 301 STGVHSCVPVGTFLQDGTGKTDIDECQTRNGGCHFCNKIVGSPDCGCKGFKLLTDEK 360  
 DB 301 STGVHSCVPVGTFLQDGTGKTDIDECQTRNGGCHFCNKIVGSPDCGCKGFKLLTDEK 360

QY 361 SCQDVDECSLDRTCDHSCINHPGTAFACNRYGTYLGFTHCGDTNECSINNGGCGQCVN 420  
 DB 361 SCQDVDECSLDRTCDHSCINHPGTAFACNRYGTYLGFTHCGDTNECSINNGGCGQCVN 420

QY 421 TVGSYECQCHPGYKLVHNNKDCVEVWGLLPTSVPRVSLHCKSGGGCGCFLRCHSG-HL 480  
 DB 421 TVGSYECQCHPGYKLVHNNKDCVEVWGLLPTSVPRVSLHCKSGGGCGCFLRCHSG-HL 480

QY 481 SSDVTTIRTSVTFKLNKGCULNAELFPEGRLPALPEKHSVSKSFVNLTCSSGQV 540  
 DB 481 SSDVTTIRTSVTFKLNKGCULNAELFPEGRLPALPEKHSVSKSFVNLTCSSGQV 540

QY 541 PGAPRSTPKEMFIVBELETNOKEVTASCDLSICVAPTEKRLKA-RTLRKAVHREQ 600  
 DB 541 PGAPRSTPKEMFIVBELETNOKEVTASCDLSICVAPTEKRLKA-RTLRKAVHREQ 600

QY 601 FHLQLSGMNLDVAKKPPRTSERQAESCGVGQGHAEHQCVSCRAGTYDGGARERCILCPNG 660  
 DB 601 FHLQLSGMNLDVAKKPPRTSERQAESCGVGQGHAEHQCVSCRAGTYDGGARERCILCPNG 660

QY 661 TFQNEBQMTCEPCPRFGNSGAJKTPEANMSECGGLCOFGSEYSAAGFAPQCCALGTQ 720  
 DB 661 TFQNEBQMTCEPCPRFGNSGAJKTPEANMSECGGLCOFGSEYSAAGFAPQCCALGTQ 720

QY 721 PEAGRTSCFPCCGGLATKHGATSCFCCETRVOCSPGHFVNTTTHRCICPVGTYQPEEG 780  
 DB 721 PEAGRTSCFPCCGGLATKHGATSCFCCETRVOCSPGHFVNTTTHRCICPVGTYQPEEG 780

QY 781 KNCVSCPGNTTTFDGSNTITCKNRRCGGELGDTFGVIESPNYPNYPANTECTWTIN 840  
 DB 781 KNCVSCPGNTTTFDGSNTITCKNRRCGGELGDTFGVIESPNYPNYPANTECTWTIN 840

QY 841 PPPKRRILLVWPEFLPIEDDCQYLVMRKTSNNSVTTETCQTYERDIATFSRSKILW 900  
 DB 841 PPPKRRILLVWPEFLPIEDDCQYLVMRKTSNNSVTTETCQTYERDIATFSRSKILW 900

QY 901 CQKSNEGNSARGFQVYVYTYDEYDELIEDIVRQGLVASENHQELKQKJLKALFDV 960  
 DB 901 CQKSNEGNSARGFQVYVYTYDEYDELIEDIVRQGLVASENHQELKQKJLKALFDV 960

QY 961 LAHPQYKXYTAGESRMPFRSFRILRSKVSFLRPYK 999  
 DB 961 LAHPQYKXYTAGESRMPFRSFRILRSKVSFLRPYK 999

RESULT 2  
 ABC05567  
 ID ABC05567 standard; Protein: 999 AA.  
 XX AC ABJ05567;  
 XX DT 14-NOV-2002 (first entry)  
 XX DE Breast cancer-associated protein 32.  
 XX KW Breast cancer; breast cancer-associated gene sequence;  
 XX KW drug development; pharmacogenetics; biosensor development.  
 XX OS Unidentified.  
 XX FN W0200259377-A2.  
 XX PD 01-AUG-2002.  
 XX PF 24-JAN-2002; 2002WO-US02242.  
 XX PR 24-JAN-2001; 2001US-263965P.  
 XX PR 02-FEB-2001; 2001US-265928P.  
 XX PR 09-APR-2001; 2001US-282698P.  
 XX PR 09-APR-2001; 2001US-0829472.  
 XX PR 04-MAY-2001; 2001US-288590P.  
 XX PR 29-MAY-2001; 2001US-294443P.  
 XX (ECGB-) EOS BIOTECHNOLOGY INC.  
 XX PA Yack DH, Gish KC, Afar D;  
 XX PI WPI; 2002-583738/62.  
 XX DR N-PSDB; ABT07724.  
 XX  
 PT Detecting a breast cancer-associated transcript in a patient's cell.  
 PT useful for diagnosing breast cancer, comprises contacting a biological  
 PT sample with a polynucleotide that selectively hybridizes with breast  
 PT cancer nucleic acids -  
 XX

PS Disclosure; Page 375; 414pp; English.

XX The invention comprises a method of detecting a breast cancer-associated transcript in a cell from a patient. The method of the invention involves contacting a biological sample from the patient with a nucleotide that hybridises to one of the 69 breast cancer-associated gene sequences shown in the specification. The method of the invention is useful in the diagnosis or prognosis of breast cancer, and for detecting genes that are up or down-regulated in breast cancer cells. Genes identified by the method of the invention can be used in diagnostic purposes and also as targets for screening for therapeutic compounds that modulate breast cancer (e.g. hormones or antibodies). Identification of genes that are over or under expressed in breast cancer can additionally provide high-resolution, high-sensitivity datasets which can be used in the areas of diagnostics, therapeutics, drug development, pharmacogenetics, protein structure and biosensor development. Amino acid sequences ABJ0536 - ABJ05604 represent the proteins encoded by the 69 breast cancer-associated genes of the invention.

XX Sequence 999 AA;

SQ

Query Match 100.0%; Score 5605; DB 23; Length 999;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVAAGRRGAAWAVLLLLPPLLLLAGAVPPGGRAGAGPQEDVDECAQGLDDCHADA 60

DB 1 MGVAAGRRGAAWAVLLLLPPLLLLAGAVPPGGRAGAGPQEDVDECAQGLDDCHADA 60

QY 61 LCQNTPTSKYCKPCYQGEQCEDECEGELNGGCVHDCNLIPGNVRCCTCPGFMIA 120

DB 61 LCQNTPTSKYCKPCYQGEQCEDECEGELNGGCVHDCNLIPGNVRCCTCPGFMIA 120

QY 121 HGHNCNLDVDECLNNGGCGQHTCVNVMGSEYCCCKEGFELSDNQHTCIHRSEGLSCMNK 180

DB 121 HGHNCNLDVDECLNNGGCGQHTCVNVMGSEYCCCKEGFELSDNQHTCIHRSEGLSCMNK 180

QY 181 DHGCSHICKEAPRGSVACRCRGFELAKNQRDCLTCNHNNGGCGQHSDDTDAGPECSC 240

DB 181 DHGCSHICKEAPRGSVACRCRGFELAKNQRDCLTCNHNNGGCGQHSDDTDAGPECSC 240

QY 241 PQYKHTDGRSCLEEDTVEVETSTSVVGDGKEVKERLLMETCAVNNGGCDRTCKDT 300

DB 241 PQYKHTDGRSCLEEDTVEVETSTSVVGDGKEVKERLLMETCAVNNGGCDRTCKDT 300

QY 301 STGVHSCPVGFTQLDGGTKCDIDECQTRNGGCDHFCNKIVGSEDCGCKGFKLLTDEK 360

DB 301 STGVHSCPVGFTQLDGGTKCDIDECQTRNGGCDHFCNKIVGSEDCGCKGFKLLTDEK 360

QY 361 SCQVDECSLDRTCDHSCINHPGTACACNRGYLYGFTHCQDTNECSINNGGCGQVCV 420

DB 361 SCQVDECSLDRTCDHSCINHPGTACACNRGYLYGFTHCQDTNECSINNGGCGQVCV 420

QY 421 TVGSYECQCHPGYKJLHNNKDCVEVKGLLPTSVSPRVSLHCGKSGGDCFLRCHSGIHL 480

DB 421 TVGSYECQCHPGYKJLHNNKDCVEVKGLLPTSVSPRVSLHCGKSGGDCFLRCHSGIHL 480

QY 481 SDDVTIIRTSVTFKLNKGLKNAELFPEGLRPALEKHSVSKESFRVNLTCSSGQV 540

DB 481 SDDVTIIRTSVTFKLNKGLKNAELFPEGLRPALEKHSVSKESFRVNLTCSSGQV 540

QY 541 PGAPGRSPKPNFIVFELETNQEVKTASDLSICIVKTEKRLKAKIRTURKAVHREQ 600

DB 541 PGAPGRSPKPNFIVFELETNQEVKTASDLSICIVKTEKRLKAKIRTURKAVHREQ 600

QY 601 FHLQSGMMLDVAKKPPRTSROAESCGVQGGHAENQCVSCRAGTYDGCARECILLCPNG 660

DB 601 FHLQSGMMLDVAKKPPRTSROAESCGVQGGHAENQCVSCRAGTYDGCARECILLCPNG 660

QY 661 TFQNEEGQMTCEPCPPRGNSGALKTPEAWNMSECGGLCPGYSADGAPACOLCALGTFQ 720

DB 661 TFQNEEGQMTCEPCPPRGNSGALKTPEAWNMSECGGLCPGYSADGAPACOLCALGTFQ 720

QY 721 PEAGRTSCFPCCGGLATKQAGATSFQCCETRVQCSQGHFVNTTTHRCIRCPVGTQPEFG 780

DB 721 PEAGRTSCFPCCGGLATKQAGATSFQCCETRVQCSQGHFVNTTTHRCIRCPVGTQPEFG 780

QY 781 KNNVCSPGNTTDFGSGTNTTCKNRRCGGEJGDFGTGYESPNYPGNYPANTECTWTIN 840

DB 781 KNNVCSPGNTTDFGSGTNTTCKNRRCGGEJGDFGTGYESPNYPGNYPANTECTWTIN 840

QY 841 PPKRRILLIVPRIFLPIEDDCGYLVWRKTSNNSVTTTCTYERPAFTSRSKKLW 900

DB 841 PPKRRILLIVPRIFLPIEDDCGYLVWRKTSNNSVTTTCTYERPAFTSRSKKLW 900

QY 901 IQPKSNEGNSARGQFQVYVYDGYGELIEDIVDRGLYASENHQELDKKLIKALFDV 960

DB 901 IQPKSNEGNSARGQFQVYVYDGYGELIEDIVDRGLYASENHQELDKKLIKALFDV 960

QY 961 LAHPONYFKYTAQESREYFPRSPFRLLRSKVSFLRPYK 999

DB 961 LAHPONYFKYTAQESREYFPRSPFRLLRSKVSFLRPYK 999

RESULT 3

ASB80926

ID ASB80926 standard; Protein: 999 AA.

XX ASB80926;

XX AC

DT 08-OCT-2002 (first entry)

XX Human breast cancer modulating protein BCO2.

DE

XX Breast cancer modulating protein; BCMP; BCO2; cytostatic; human;

KW immunostimulant; antisense therapy; gene therapy; vaccine.

XX Homo sapiens.

XX WO200255988-A2.

XX 18-JUL-2002.

PF 21-DEC-2001; 2001WO-US50548.

XX 21-DEC-2000; 2000US-0747371.

XX (ECSB-); EOS BIOTECHNOLOGY INC.

PA Gish KC, Mack D;

XX WPT; 2002-566749/60.

XX N-PSDB; ARN86363.

XX Diagnosing breast cancer comprises comparing the gene expression of a

PT breast cancer modulating protein in a test breast tissue and unaffected

PT breast tissue of the same or a different patient

XX Disclosure; Fig 2; 68pp; English.

XX The invention relates to diagnosing breast cancer that involves

CC determining the expression of a gene encoding breast cancer modulating

CC protein (BCO2) or a fragment of it in a breast tissue of an individual,

CC and comparing the expression of the genes from a second normal breast

CC tissue from the individual or a second unaffected individual, where a

CC difference in expression indicates the individual has breast cancer. The

CC method is used for diagnosing breast cancer. Other new methods are

CC provided for screening drug candidates, screening for bioactive agents,

CC and evaluating the effect of a candidate breast cancer drug. An antibody

CC to BCO2 can be used to inhibit the bioactivity of BCO2 or neutralise it's

CC effect and can be used to treat breast cancer. The antibody can also be

CC used to localise a therapeutic group to breast cancer tissue, which can

CC be used to treat breast cancer. Antisense molecules are used to inhibit

CC breast cancer in a cell. BCO2 protein and polynucleotides are used to

CC elicit an immune response in an individual. BCO2 is also used to

CC determining the prognosis of an individual with breast cancer. The

CC present sequence represents the human BCO2 protein.

XX	Sequence	999 AA;	Query Match	100.0%;	Score 5605;	DB 23;	Length 999;
XX	Best Local Similarity	100.0%;	Pred. No. 0;				
XX	Matches 999;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
Qy	1	MGVAGNRPGAAWAVLLLLLLLLLLLLLAGAVPPGGRAGAPQEDVDECAQGLDCHADA	60				
Dy	1	MGVAGNRPGAAWAVLLLLLLLLLLLLLAGAVPPGGRAGAPQEDVDECAQGLDCHADA	60				
Qy	61	LCQNTPTSYKSCCKPGVGGGROCEDIDECGNELNGSCVHDCUNIPGNVRCCTCFGEVLA	120				
Dy	61	LCQNTPTSYKSCCKPGVGGGROCEDIDECGNELNGSCVHDCUNIPGNVRCCTCFGEVLA	120				
Qy	121	HDGHNCLDVECLJENNGGQHTCVNMGSEYECCKEGFFLSNQHTCIHRSSEGLSCNMK	180				
Dy	121	HDGHNCLDVECLJENNGGQHTCVNMGSEYECCKEGFFLSNQHTCIHRSSEGLSCNMK	180				
Qy	181	DHGCISHICREAPRGVACBPCPGFELAKNQRDCILTCNHNNGGCGQSCDDTADGPECSCH	240				
Dy	181	DHGCISHICREAPRGVACBPCPGFELAKNQRDCILTCNHNNGGCGQSCDDTADGPECSCH	240				
Qy	241	POYKMTDGRSLEREDTVLEVTESNTTSVVDGKPKVRRLLMETCAVNGGCGDRTCKDT	300				
Dy	241	POYKMTDGRSLEREDTVLEVTESNTTSVVDGKPKVRRLLMETCAVNGGCGDRTCKDT	300				
Qy	301	STGVHSCVPVFTLQDGTCKDIDECQTRNGGCDHFCKNIYVGFPCGCKGFKLLTDEK	360				
Dy	301	STGVHSCVPVFTLQDGTCKDIDECQTRNGGCDHFCKNIYVGFPCGCKGFKLLTDEK	360				
Qy	361	SCQDVDECSLDRTCDHSCINHGRTFACACNRYTLVGFTHCGDTNRCSTNNGCQGVN	420				
Dy	361	SCQDVDECSLDRTCDHSCINHGRTFACACNRYTLVGFTHCGDTNRCSTNNGCQGVN	420				
Qy	421	TVGSYECQCHPGYKJLHNNKDKVEVKGLLPTSVSRVSLHCGKGGGCGFLRCHSGIHL	480				
Dy	421	TVGSYECQCHPGYKJLHNNKDKVEVKGLLPTSVSRVSLHCGKGGGCGFLRCHSGIHL	480				
Qy	481	SSDVTIRTSTVKLNGKGLSKNAELFPEGLRALPEKHSVVKESFRVYNTTCSGKGV	540				
Dy	481	SSDVTIRTSTVKLNGKGLSKNAELFPEGLRALPEKHSVVKESFRVYNTTCSGKGV	540				
Qy	541	PGAPGRPSTPKEMFTTVEPELETNQEVNTASCDLSICIVKRTKRLKARTLRKAVHREQ	600				
Dy	541	PGAPGRPSTPKEMFTTVEPELETNQEVNTASCDLSICIVKRTKRLKARTLRKAVHREQ	600				
Qy	601	FHLQSGMNLDAVKPPTTSERQAESGVGGQGHAEHQCVSCRAGTYVYGARERCILCPNG	660				
Dy	601	FHLQSGMNLDAVKPPTTSERQAESGVGGQGHAEHQCVSCRAGTYVYGARERCILCPNG	660				
Qy	661	TFONEGOMTCEPCPPGNSGALKTPKANMSGCGLCOPGEVSADGAPCOLCALGTFQ	720				
Dy	661	TFONEGOMTCEPCPPGNSGALKTPKANMSGCGLCOPGEVSADGAPCOLCALGTFQ	720				
Qy	721	PEAGRTSCFFCGGLATKHGATSFQDCETRVQCSQGHFVNTTTHRCIRCPVGTQPEFG	780				
Dy	721	PEAGRTSCFFCGGLATKHGATSFQDCETRVQCSQGHFVNTTTHRCIRCPVGTQPEFG	780				
Qy	781	KNNVCSPGNTTDFDGSNTITCKNRRCGGELGDFGTGIESPNYPGNTPANTECTWTN	840				
Dy	781	KNNVCSPGNTTDFDGSNTITCKNRRCGGELGDFGTGIESPNYPGNTPANTECTWTN	840				
Qy	841	P2PKRRLIVVPELFLP1EDDCDYLVMKRTSSNSUTTYETCTQYERPIATSRSKLM	900				
Dy	841	P2PKRRLIVVPELFLP1EDDCDYLVMKRTSSNSUTTYETCTQYERPIATSRSKLM	900				
Qy	901	IQFKNSEGNARGQVPPVYTYDEYQELIEDIVRDGRLVASENHQBILKDKKLIKALFV	960				
Dy	901	IQFKNSEGNARGQVPPVYTYDEYQELIEDIVRDGRLVASENHQBILKDKKLIKALFV	960				
Qy	961	LAHPQNYFKYTAQESREMPFRSFIIRLLRSKVSRLRPAK 999					

Db 961 LAHPQNYFKYTAQESREMPFRSFIIRLLRSKVSRLRPAK 999

RESULT 4  
 ABG61893  
 ID ABG61893 standard; Protein; 999 AA.  
 XX  
 AC ABG61893;  
 XX  
 DT 15-AUG-2002 (first entry)  
 XX  
 DE Prostate cancer-associated protein #94.  
 XX  
 KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.  
 XX  
 OS Mammalia.  
 PN WO200230268-A2.  
 XX  
 PD 19-APR-2002.  
 XX  
 PF 12-OCT-2001; 2001WC-US32045.  
 XX  
 PR 13-OCT-2000; 2000US-0687576.  
 PR 08-DEC-2000; 2000US-0733388.  
 ER 08-DEC-2000; 2000US-0733342.  
 PR 24-JAN-2001; 2001US-263957P.  
 PR 16-MAR-2001; 2001US-276797P.  
 PR 06-APR-2001; 2001US-261922P.  
 PR 24-APR-2001; 2001US-286214P.  
 PR 30-APR-2001; 2001US-0847046.  
 PR 04-MAY-2001; 2001US-286599P.  
 XX  
 PA (EOSB-) EOS BIOTECHNOLOGY INC.  
 XX  
 PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;  
 XX  
 DR WPI; 2002-471335/50.  
 DR N-PSDB; ABK92209.  
 XX  
 DT Detecting a prostate cancer-associated transcript in a cell in a patient, useful for diagnosing prostate cancer (PC) or screening modulators of PC, by determining if prostate cancer-associated genes are expressed in a prostate tissue  
 XX  
 PS Claim 27; Page 379; 436pp; English.  
 XX  
 CC The present invention relates to methods of detecting a prostate cancer-associated transcript in a cell from a patient. The method comprises contacting a biological sample from the patient with prostate cancer-associated polynucleotides (designated PC genes) that selectively hybridise to a sequence that is at least 80% identical to them. The prostate cancer-associated polynucleotide sequences are differentially expressed in prostate tumour tissue or in prostate cancer and are derived from the tissues of various organisms such as humans or other mammals (e.g. mice, sheep and dogs). The methods of the invention are useful for diagnosing and treating prostate cancer in mammals. The prostate cancer-associated genes are useful for diagnosing or treating prostate cancer, as well as for identifying modulators of prostate cancer or agents that inhibit prostate cancer. The nucleic acid sequences are particularly useful in gene therapy, as a vaccine or in antisense applications.  
 CC ABG61800-ABG61944 represent prostate cancer-associated proteins.  
 XX  
 SQ Sequence 999 AA;

Query Match 100.0%; Score 5605; DB 23; Length 999;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy : MGVAGNRPGAAWAVLLLLLLLLLAGAVPPGGRAGAPQEDVDECAQGLDCHADA 60

```

Db      1  MGVAGRNRPGAAWAVLLLLPPLLLLAGAVPPGRGAAAGFQEDVDECAQGLDDCHADA 60
Qy      61  LCQNTPTSYKCKPGYQGEGRQCEDIDECGNELNGGCVHDCNLNPGNVRCTCFDGFMLA 120
Db      61  LCQNTPTSYKCKPGYQGEGRQCEDIDECGNELNGGCVHDCNLNPGNVRCTCFDGFMLA 120
Qy      121  HDGNCILDVDECLNNGGCGQHTCVNVMGSYECCKEGFFLSDNQHTCIHRSEGLSCNKK 180
Db      121  HDGNCILDVDECLNNGGCGQHTCVNVMGSYECCKEGFFLSDNQHTCIHRSEGLSCNKK 180
Qy      181  DHGCSHICKEAPRGSVACRCRPGFELAKNCRCLITCNHNGGCGCHSCDDTADGECSC 240
Db      181  DHGCSHICKEAPRGSVACRCRPGFELAKNCRCLITCNHNGGCGCHSCDDTADGECSC 240
Qy      241  POYKMTDGRSCLEREDTVLEVTESNTTSVWDGDKRVKRLMETCAVNNGGCDRTCKDT 300
Db      241  POYKMTDGRSCLEREDTVLEVTESNTTSVWDGDKRVKRLMETCAVNNGGCDRTCKDT 300
Qy      301  STGVHSCVPVFTLQLOGTKCDIDECQTRNGGCHFCCKNI VGSFDCGCKGFKLLTDEK 360
Db      301  STGVHSCVPVFTLQLOGTKCDIDECQTRNGGCHFCCKNI VGSFDCGCKGFKLLTDEK 360
Qy      361  SCQDVDECSLDRTCDHSCINHPCTTACACNRGVTLYGFTHCQDTNECSINNGGCGOVCVN 420
Db      361  SCQDVDECSLDRTCDHSCINHPCTTACACNRGVTLYGFTHCQDTNECSINNGGCGOVCVN 420
Qy      421  TVGSYECQCHPGYKHLHNKKDCVEYKGLLPTSVSRVSLHCGKSGGDCGCFLRCHSGIHL 480
Db      421  TVGSYECQCHPGYKHLHNKKDCVEYKGLLPTSVSRVSLHCGKSGGDCGCFLRCHSGIHL 480
Qy      481  SSVVTTIRTSVTFKLNKGKSLKNAELFPEGLRPALPEKHSSVKESFRYVNLTCSSGQV 540
Db      481  SSVVTTIRTSVTFKLNKGKSLKNAELFPEGLRPALPEKHSSVKESFRYVNLTCSSGQV 540
Qy      541  PGAPGRPSTPKMFIIVFELETNQKENTASDLSQIVKRTKRLKAIKRTLRKAVHREQ 600
Db      541  PGAPGRPSTPKMFIIVFELETNQKENTASDLSQIVKRTKRLKAIKRTLRKAVHREQ 600
Qy      601  FHLQSGMNLVAKKPPRTSRAEQSCVGGQHAENQCVSCRAGTYDGARERCLCPNG 660
Db      601  FHLQSGMNLVAKKPPRTSRAEQSCVGGQHAENQCVSCRAGTYDGARERCLCPNG 660
Qy      661  TQNEBQMTCEPCPPNGSGALKTPAANNSECGGLQCPGYSADGFAPCOLCAJTFQ 720
Db      661  TQNEBQMTCEPCPPNGSGALKTPAANNSECGGLQCPGYSADGFAPCOLCAJTFQ 720
Qy      721  PEAGRTSCFPCCGGLATKHGATSFQDCETRVQCSGPHFYNTTTHRCRCPVGTQPEFG 780
Db      721  PEAGRTSCFPCCGGLATKHGATSFQDCETRVQCSGPHFYNTTTHRCRCPVGTQPEFG 780
Qy      791  KNCVSCPGNTTDFDGSNTITQCKNRRCGELGDFTVIESPNYPGNYPANTCTWTIN 840
Db      781  KNCVSCPGNTTDFDGSNTITQCKNRRCGELGDFTVIESPNYPGNYPANTCTWTIN 840
Qy      841  PPKRRILLVWEIFLPIEDDCGYLWVRKTSSSNSVTTFYTCQYERPIATSKSLW 900
Db      841  PPKRRILLVWEIFLPIEDDCGYLWVRKTSSSNSVTTFYTCQYERPIATSKSLW 900
Qy      901  IQFKSNEGNSARGFQVYVYDYDEYQELIEDIVRGRLYASENHQELIKKLLKALPDV 960
Db      901  IQFKSNEGNSARGFQVYVYDYDEYQELIEDIVRGRLYASENHQELIKKLLKALPDV 960
Qy      961  LAHPQNYKYTAQESREMPFRSIRLLRSKVRFLAPYK 999
Db      961  LAHPQNYKYTAQESREMPFRSIRLLRSKVRFLAPYK 999

RESULT 5
ABR47407
ID ABR47407 standard; Protein: 999 AA.
XX
AC ABR47407;
```

```

XX      12-JUN-2003 (first entry)
XX      Breast cancer associated protein sequence SEQ ID NO:45.
XX      Human; breast cancer; cytostatic; gene therapy.
XX      Homo sapiens.
XX      WO2003004999-A2.
XX      16-JAN-2003.
XX      21-JUN-2002; 2002WO-US19669.
XX      21-JUN-2001; 2001US-299887P.
XX      27-JUN-2001; 2001US-301572P.
XX      18-JUL-2001; 2001US-306501P.
XX      25-SEP-2001; 2001US-325002P.
XX      05-MAR-2002; 2002US-362585P.
XX      14-MAY-2002; 2002US-380391P.
XX      (MILL-) MILLENIUM PHARM INC.
XX      Lillie J, Gannavarapu M, Glatt K, Hoersch S, Kamatkar S, Mertens M,
XX      Monahan JS, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE, East RC,
XX      Hortobagyi GN, Fuzsai L, Meric F, Sahin A, Mills GB;
XX      WPI; 2003-2:0381/20.
XX      N-PDSB; ACC50098.
XX      Breast cancer diagnosis or treatment by comparing the level of
XX      expression of a marker in a patient sample with that in the control;
XX      non-breast cancer sample.
XX      Claim 1; SEQ ID 45; 128pp; English.
XX      The present invention describes a method for assessing whether a patient
XX      is afflicted with breast cancer. The method comprises comparing the level
XX      of expression of a marker (gene/protein) to ACC50076 to ACC50334 and
XX      ABR47365 to ABR47632) in a patient sample and the normal level of
XX      expression of the marker in a control non-breast cancer sample, where a
XX      significant increase in the level of expression of the marker in the
XX      patient sample and the normal level is an indication that the patient is
XX      afflicted with breast cancer. The breast cancer associated sequences
XX      from the present invention have cytostatic activities and can be used in
XX      gene therapy. The method is useful for diagnosing and treating breast
XX      cancer.
XX      N.B. The sequence data for this patent did not form part of the printed
XX      specification, but was obtained in electronic format directly from WIPO
XX      at ftp.wipo.int/pub/published_pat_sequences.
XX      Sequence 999 AA;
XX      Query March 100.0%; Score 5605; DB 24; Length 999;
XX      Best Local Similarity 100.0%; Pred. No. 0;
XX      Matches 999; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX      1  MGVAGRNRPGAAWAVLLLLPPLLLLAGAVPPGRGAAAGFQEDVDECAQGLDDCHADA 60
XX      1  MGVAGRNRPGAAWAVLLLLPPLLLLAGAVPPGRGAAAGFQEDVDECAQGLDDCHADA 60
XX      61  LCQNTPTSYKCKPGYQGEGRQCEDIDECGNELNGGCVHDCNLNPGNVRCTCFDGFMLA 120
XX      61  LCQNTPTSYKCKPGYQGEGRQCEDIDECGNELNGGCVHDCNLNPGNVRCTCFDGFMLA 120
XX      121  HDGNCILDVDECLNNGGCGQHTCVNVMGSYECCKEGFFLSDNQHTCIHRSEGLSCNKK 180
XX      121  HDGNCILDVDECLNNGGCGQHTCVNVMGSYECCKEGFFLSDNQHTCIHRSEGLSCNKK 180
XX      181  DHGCSHICKEAPRGSVACRCRPGFELAKNCRCLITCNHNGGCGCHSCDDTADGECSC 240
XX      181  DHGCSHICKEAPRGSVACRCRPGFELAKNCRCLITCNHNGGCGCHSCDDTADGECSC 240
XX      241  POYKMTDGRSCLEREDTVLEVTESNTTSVWDGDKRVKRLMETCAVNNGGCDRTCKDT 300
XX      241  POYKMTDGRSCLEREDTVLEVTESNTTSVWDGDKRVKRLMETCAVNNGGCDRTCKDT 300
XX      301  STGVHSCVPVFTLQLOGTKCDIDECQTRNGGCHFCCKNI VGSFDCGCKGFKLLTDEK 360
XX      301  STGVHSCVPVFTLQLOGTKCDIDECQTRNGGCHFCCKNI VGSFDCGCKGFKLLTDEK 360
XX      361  SCQDVDECSLDRTCDHSCINHPCTTACACNRGVTLYGFTHCQDTNECSINNGGCGOVCVN 420
XX      361  SCQDVDECSLDRTCDHSCINHPCTTACACNRGVTLYGFTHCQDTNECSINNGGCGOVCVN 420
XX      421  TVGSYECQCHPGYKHLHNKKDCVEYKGLLPTSVSRVSLHCGKSGGDCGCFLRCHSGIHL 480
XX      421  TVGSYECQCHPGYKHLHNKKDCVEYKGLLPTSVSRVSLHCGKSGGDCGCFLRCHSGIHL 480
XX      481  SSVVTTIRTSVTFKLNKGKSLKNAELFPEGLRPALPEKHSSVKESFRYVNLTCSSGQV 540
XX      481  SSVVTTIRTSVTFKLNKGKSLKNAELFPEGLRPALPEKHSSVKESFRYVNLTCSSGQV 540
XX      541  PGAPGRPSTPKMFIIVFELETNQKENTASDLSQIVKRTKRLKAIKRTLRKAVHREQ 600
XX      541  PGAPGRPSTPKMFIIVFELETNQKENTASDLSQIVKRTKRLKAIKRTLRKAVHREQ 600
XX      601  FHLQSGMNLVAKKPPRTSRAEQSCVGGQHAENQCVSCRAGTYDGARERCLCPNG 660
XX      601  FHLQSGMNLVAKKPPRTSRAEQSCVGGQHAENQCVSCRAGTYDGARERCLCPNG 660
XX      661  TQNEBQMTCEPCPPNGSGALKTPAANNSECGGLQCPGYSADGFAPCOLCAJTFQ 720
XX      661  TQNEBQMTCEPCPPNGSGALKTPAANNSECGGLQCPGYSADGFAPCOLCAJTFQ 720
XX      721  PEAGRTSCFPCCGGLATKHGATSFQDCETRVQCSGPHFYNTTTHRCRCPVGTQPEFG 780
XX      721  PEAGRTSCFPCCGGLATKHGATSFQDCETRVQCSGPHFYNTTTHRCRCPVGTQPEFG 780
XX      791  KNCVSCPGNTTDFDGSNTITQCKNRRCGELGDFTVIESPNYPGNYPANTCTWTIN 840
XX      781  KNCVSCPGNTTDFDGSNTITQCKNRRCGELGDFTVIESPNYPGNYPANTCTWTIN 840
XX      841  PPKRRILLVWEIFLPIEDDCGYLWVRKTSSSNSVTTFYTCQYERPIATSKSLW 900
XX      841  PPKRRILLVWEIFLPIEDDCGYLWVRKTSSSNSVTTFYTCQYERPIATSKSLW 900
XX      901  IQFKSNEGNSARGFQVYVYDYDEYQELIEDIVRGRLYASENHQELIKKLLKALPDV 960
XX      901  IQFKSNEGNSARGFQVYVYDYDEYQELIEDIVRGRLYASENHQELIKKLLKALPDV 960
XX      961  LAHPQNYKYTAQESREMPFRSIRLLRSKVRFLAPYK 999
XX      961  LAHPQNYKYTAQESREMPFRSIRLLRSKVRFLAPYK 999
```



QY 241 POYKMTDGRSCLEREDTVLEVTESNTTSVVDGDKSVKRLMETCAVNNGGCDRTCKDT 300  
 DB 241 POYKMTDGRSCLEREDTVLEVTESNTTSVVDGDKSVKRLMETCAVNNGGCDRTCKDT 300  
 QY 301 STGVHCSCPVGTFLQDGTCKDIDECOTRNGGCDHFCNKIVGSDCCCKGKGFLLTDEK 360  
 DB 301 STGVHCSCPVGTFLQDGTCKDIDECOTRNGGCDHFCNKIVGSDCCCKGKGFLLTDEK 360  
 QY 361 SCQVDECSLDRTCDHSCINHPGTTFACACNRGYTLVGFTHCGDTNECSINNGCQVCVN 420  
 DB 361 SCQVDECSLDRTCDHSCINHPGTTFACACNRGYTLVGFTHCGDTNECSINNGCQVCVN 420  
 QY 421 TVGSVECOCHPGYKLVHMKDCEVKGLLPTSVSRVSLHCKSGGGCGCFRCHSGIHL 480  
 DB 421 TVGSVECOCHPGYKLVHMKDCEVKGLLPTSVSRVSLHCKSGGGCGCFRCHSGIHL 480  
 QY 481 SSVDTTIRTSVTFKLNKCSLKNALPEGLRPALEKHSVKESFRVNLTCSSGQV 540  
 DB 481 SSVDTTIRTSVTFKLNKCSLKNALPEGLRPALEKHSVKESFRVNLTCSSGQV 540  
 QY 541 PGAPGRSPCKEMFTVEPELTNKEVTASCDLSICVKTETKRLKAIKRLKAVHREQ 600  
 DB 541 PGAPGRSPCKEMFTVEPELTNKEVTASCDLSICVKTETKRLKAIKRLKAVHREQ 600  
 QY 601 FHLQSGMNLVAKPPRTSERQAESCGVQGHAEHQVSCBAGTYDGBARCLCPNG 660  
 DB 601 FHLQSGMNLVAKPPRTSERQAESCGVQGHAEHQVSCBAGTYDGBARCLCPNG 660  
 QY 661 TFCNEGOMTCPCPRPGNSGALKTPKPEANMSECGGLQCPGEYSADGAPCOLCALGTFO 720  
 DB 661 TFCNEGOMTCPCPRPGNSGALKTPKPEANMSECGGLQCPGEYSADGAPCOLCALGTFO 720  
 QY 721 PEAGRTSCPPCGGLATKHQATSFQDCETRVQCSPGHFYNTTTHRCRCVPVGTQPEFG 780  
 DB 721 PEAGRTSCPPCGGLATKHQATSFQDCETRVQCSPGHFYNTTTHRCRCVPVGTQPEFG 780  
 QY 781 KNNVSCPGNTTTFDGSNTITQCKNRCCGGLGFTGVIESPNYPNPANTECTWTIN 840  
 DB 781 KNNVSCPGNTTTFDGSNTITQCKNRCCGGLGFTGVIESPNYPNPANTECTWTIN 840  
 QY 841 PPPKRLIIVWPEIFLPIEDDCGDLVWPKTSNNSVTTTETCOTYERPIAFTSRSKLM 900  
 DB 841 PPPKRLIIVWPEIFLPIEDDCGDLVWPKTSNNSVTTTETCOTYERPIAFTSRSKLM 900  
 QY 901 IOFKNEGNSARGFQVYVYDDEYQELIEDIVROGRLYASENHQELIKKLIKALFDV 960  
 DB 901 IOFKNEGNSARGFQVYVYDDEYQELIEDIVROGRLYASENHQELIKKLIKALFDV 960  
 QY 961 LAHPQNYFYTAQEGREMPFRFIRLLRSKVSFLRPYK 999  
 DB 961 LAHPQNYFYTAQEGREMPFRFIRLLRSKVSFLRPYK 999

RESULT 6  
 ID ABJ19815  
 XX ABJ19815 standard; Protein; 999 AA.  
 AC ABJ19815;  
 XX  
 DT 10-APR-2003 (first entry)  
 XX  
 DE Androgen-independent prostate cancer-related protein - SEQ ID NO 42.  
 XX Androgen-independent cancer; androgen ablation therapy; prostate cancer;  
 KW androgen-dependent prostate cancer; prostate cancer.  
 XX  
 CS Unidentified.  
 XX  
 FN WO200298358-A2.  
 XX  
 PD 12-DEC-2002.  
 XX

PF 04-JUN-2002; 2002WO-US17594.  
 XX 04-JUN-2001; 2001US-2959:7P.  
 PR 13-NOV-2001; 2001US-353666P.  
 PR 29-MAR-2002; 2002US-368689P.  
 PR 12-APR-2002; 2002US-372246P.  
 PR 31-MAY-2002; 2002US-0160233.  
 XX (ECSE-) BOS BIOTECHNOLOGY INC.  
 XX Afar DEH, Agus C, Mack DH;  
 XX WPI; 2003-148602/24.  
 PT Detecting an androgen-independent prostate cancer cell in a sample or  
 PT diagnosing androgen-dependent prostate cancer, by determining the  
 PT presence or absence of genes whose expressions are up- or  
 PT down-regulated.  
 XX  
 PS Claim 1; Page 199; 210pp; English.  
 XX  
 CC The invention comprises a method for detecting an androgen-independent  
 CC cancer cell in a sample from a patient who has undergone androgen  
 CC ablation therapy. The method involves determining the presence or absence  
 CC of nucleic acids that are either up-regulated or down-regulated in  
 CC prostate cancer. The method is useful for detecting an androgen-  
 CC independent prostate cancer cell in a sample from a patient who has  
 CC undergone androgen ablation therapy. The method is particularly useful  
 CC for diagnosing androgen-dependent prostate cancer, prostate cancer  
 CC undergoing androgen withdrawal, or androgen-independent prostate cancer.  
 CC The present amino acid sequence represents a protein which is encoded by  
 CC a gene that is either up-regulated or down-regulated in prostate cancer.  
 XX  
 SQ Sequence 999 AA;

Query Match: 100.0%; Score 5605; DB 24; Length 999;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 999; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MSVAGNRPGAAWAVJLLLLPPLLJAGAVPGRGAAAGPOEDVDECAQGLDCHADA 60  
 DB 1 MSVAGNRPGAAWAVJLLLLPPLLJAGAVPGRGAAAGPOEDVDECAQGLDCHADA 60  
 QY 61 LCQNTPTSYKCSKPGYQGGEGROCEIDECNGLNGGCVHDCNIPGNVRCCTCFDGMJA 120  
 DB 61 LCQNTPTSYKCSKPGYQGGEGROCEIDECNGLNGGCVHDCNIPGNVRCCTCFDGMJA 120  
 QY 121 HDGHNCLDVEDELENNGGCGQHTCVNVMGSVECCCKEGFLSDNQHTCIHRSEGLSCMK 180  
 DB 121 HDGHNCLDVEDELENNGGCGQHTCVNVMGSVECCCKEGFLSDNQHTCIHRSEGLSCMK 180  
 QY 181 DHGCHICKAPRGSVACBPCGFELAKNORDILTCNHGNSGCCSCDQTAGPSCSCH 240  
 DB 181 DHGCHICKAPRGSVACBPCGFELAKNORDILTCNHGNSGCCSCDQTAGPSCSCH 240  
 QY 241 PQYKHTDGRSCLEREDTVLEVTESNTTSVVDGDKSVKRLMETCAVNNGGCDRTCKDT 300  
 DB 241 PQYKHTDGRSCLEREDTVLEVTESNTTSVVDGDKSVKRLMETCAVNNGGCDRTCKDT 300  
 QY 301 STGVHCSCPVGTFLQDGTCKDIDECOTRNGGCDHFCNKIVGSDCCCKGKGFLLTDEK 360  
 DB 301 STGVHCSCPVGTFLQDGTCKDIDECOTRNGGCDHFCNKIVGSDCCCKGKGFLLTDEK 360  
 QY 361 SCQVDECSLDRTCDHSCINHPGTTFACACNRGYTLVGFTHCGDTNECSINNGCQVCVN 420  
 DB 361 SCQVDECSLDRTCDHSCINHPGTTFACACNRGYTLVGFTHCGDTNECSINNGCQVCVN 420  
 QY 421 TVGSVECOCHPGYKLVHMKDCEVKGLLPTSVSRVSLHCKSGGGCGCFRCHSGIHL 480  
 DB 421 TVGSVECOCHPGYKLVHMKDCEVKGLLPTSVSRVSLHCKSGGGCGCFRCHSGIHL 480  
 QY 481 SSVDTTIRTSVTFKLNKCSLKNALPEGLRPALEKHSVKESFRVNLTCSSGQV 540

Db 481 SSDVTTIRTSVTFKLNKGKSLKNAELFPEGLRALPEKHSSVKESFRYNLTCCSGKQV 540  
 Qy 541 PGAPGRPSTPKENFIVFELETNQKEVTASCDLSICIVKRTKRLKAIKRTLRKAVHREQ 600  
 Db 541 PGAPGRPSTPKENFIVFELETNQKEVTASCDLSICIVKRTKRLKAIKRTLRKAVHREQ 600  
 Qy 601 FHLQSLGMNLDVAKKPPRTSERQAESCGVGQGAENQCVSCRASTYDYGARECILCPNG 660  
 Db 601 FHLQSLGMNLDVAKKPPRTSERQAESCGVGQGAENQCVSCRASTYDYGARECILCPNG 660  
 Qy 661 TFQNEEQMTCEPCPRPGNSGALKTPAAMNMSCGGLCPGGEYSADGFAPQCALGTQ 720  
 Db 661 TFQNEEQMTCEPCPRPGNSGALKTPAAMNMSCGGLCPGGEYSADGFAPQCALGTQ 720  
 Qy 721 PEAGRTSCPCGGGLATKHGATSFQDCETRVCQSPGHFNTTTHRCIRCPVGTQPEFG 780  
 Db 721 PEAGRTSCPCGGGLATKHGATSFQDCETRVCQSPGHFNTTTHRCIRCPVGTQPEFG 780  
 Qy 781 KNNVCSCPGNTTDFDGSNTNITCKNRRCGGELGDTGYIESPNYGNYPANTECTWTIN 840  
 Db 781 KNNVCSCPGNTTDFDGSNTNITCKNRRCGGELGDTGYIESPNYGNYPANTECTWTIN 840  
 Qy 841 PPPKRRILIVVPEIFLPIEDDCDGLVMKRTSSNSVTYETCQTYERPIAFTSRSKLW 900  
 Db 841 PPPKRRILIVVPEIFLPIEDDCDGLVMKRTSSNSVTYETCQTYERPIAFTSRSKLW 900  
 Qy 901 IQPKSNEGSARGQVPPVYVYDQYQELIEDIVRDGRLYASENHQELKOKLIKALFDV 960  
 Db 901 IQPKSNEGSARGQVPPVYVYDQYQELIEDIVRDGRLYASENHQELKOKLIKALFDV 960  
 Qy 961 LAHPQNYFKYTAQESREMPFRSIRLLRSKVSRLRPYK 999  
 Db 961 LAHPQNYFKYTAQESREMPFRSIRLLRSKVSRLRPYK 999

RESULT 7

ABJ37049  
 ID ABJ37049 standard; Protein; 997 AA.  
 XX AC ABJ37049;  
 XX D7 01-MAY-2003 (first entry)  
 XX DE Human: breast cancer / ovarian cancer related protein #25.  
 XX KW Human; cytostatic; breast cancer; ovarian cancer.  
 XX OS Homo sapiens.  
 XX PN W02003000C12-A2.  
 XX PD 03-JAN-2003.  
 XX PF 21-JUN-2002; 2002WO-US19773.  
 XX PR 21-JUN-2001; 2001US-300159P.  
 XX PR 27-JUN-2001; 2001US-301351P.  
 XX (MILL-) MILLENNIUM PHARM INC.  
 XX PA Veiby OP;  
 XX PI WP2; 2003-267848/26.  
 XX DR N-PSDB; ABT31918.  
 XX PT Determining the presence of breast cancer in an individual, involves  
 XX PT using specific polynucleotide markers -  
 XX PS Disclosure; Page 141-143; 233pp; English.  
 XX CC The invention comprises a method for assessing whether a patient is  
 XX CC afflicted with breast cancer or ovarian cancer. The method involves the  
 XX CC use of specific DNA markers. The method of the invention is useful in the

CC detection and treatment of ovarian and breast cancer. Amino acid  
 CC sequences ABJ37049 - ABJ37049 represent human breast/ovarian cancer-  
 CC related proteins.  
 XX  
 SC Sequence 997 AA:  
 Query Match 99.6%; Score 5581; DB 24; Length 997;  
 Best Local Similarity 99.7%; Pred. No. 3;  
 Matches 996; Conservative 0; Mismatches 1; Indels 2; Gaps 1;  
 Qy 1 MGVAGRNRPAGAAVALLLLPPLLAGAVPPGRGAAQPOEDVDECAQGLDCHADA 60  
 Db 1 MGVAGRNRPAGAAVALLLLPPLLAGAVPPGRGAAQPOEDVDECAQGLDCHADA 58  
 Qy 61 LCQNTPTSYKCSCKPVGQGEGRQCEDIDECNENLNGCVHDCNLIPGNYRCTCDGFMLA 120  
 Db 59 LCQNTPTSYKCSCKPVGQGEGRQCEDIDECNENLNGCVHDCNLIPGNYRCTCDGFMLA 118  
 Qy 121 HDGNCILDVDCLENNGGCQHTCVNMGSYECCCKEGFFLSNCHTCIHRSEBGLSCNKK 180  
 Db 119 HDGNCILDVDCLENNGGCQHTCVNMGSYECCCKEGFFLSNCHTCIHRSEBGLSCNKK 178  
 Qy 181 DHGCSHICKEAPRGSAVACECPGFELAKNQDCILTCNHGNGGCGCHSCDDTADGPECSCH 240  
 Db 179 DHGCSHICKEAPRGSAVACECPGFELAKNQDCILTCNHGNGGCGCHSCDDTADGPECSCH 238  
 Qy 241 POYKMHIDGRSCLEREDTVLEVTESENTTSVVDGDKPVKRRLLMETCAVNNCGCDRTCKDT 300  
 Db 239 POYKMHIDGRSCLEREDTVLEVTESENTTSVVDGDKPVKRRLLMETCAVNNCGCDRTCKDT 298  
 Qy 301 STGVHSCSPVGTFLQDCKTKDIDECOTRNGGCDHFCNKIVGSPDCGCKKGFLLTDEK 360  
 Db 299 STGVHSCSPVGTFLQDCKTKDIDECOTRNGGCDHFCNKIVGSPDCGCKKGFLLTDEK 358  
 Qy 361 SCCQVDECSLDRTCDHSCINHPGTACACNRYTLYGFTHCDDTNECSINNGGCGQVQVW 420  
 Db 359 SCCQVDECSLDRTCDHSCINHPGTACACNRYTLYGFTHCDDTNECSINNGGCGQVQVW 418  
 Qy 421 TVGSYECQCHPGYKLNHKKDCVEKGLLPVSVSPRSLHCKSGGGGDCFLRCHSGIHL 490  
 Db 419 TVGSYECQCHPGYKLNHKKDCVEKGLLPVSVSPRSLHCKSGGGGDCFLRCHSGIHL 478  
 Qy 481 SSDVTTIRTSVTFKLNKGKSLKNAELFPEGLRALPEKHSSVKESFRYNLTCCSGKQV 540  
 Db 479 SSDVTTIRTSVTFKLNKGKSLKNAELFPEGLRALPEKHSSVKESFRYNLTCCSGKQV 538  
 Qy 541 PGAPGRPSTPKENFIVFELETNQKEVTASCDLSICIVKRTKRLKAIKRTLRKAVHREQ 600  
 Db 539 PGAPGRPSTPKENFIVFELETNQKEVTASCDLSICIVKRTKRLKAIKRTLRKAVHREQ 598  
 Qy 601 FHLQSLGMNLDVAKKPPRTSERQAESCGVGQGAENQCVSCRASTYDYGARECILCPNG 660  
 Db 599 FHLQSLGMNLDVAKKPPRTSERQAESCGVGQGAENQCVSCRASTYDYGARECILCPNG 658  
 Qy 661 TFQNEEQMTCEPCPRPGNSGALKTPAAMNMSCGGLCPGGEYSADGFAPQCALGTQ 720  
 Db 659 TFQNEEQMTCEPCPRPGNSGALKTPAAMNMSCGGLCPGGEYSADGFAPQCALGTQ 718  
 Qy 721 PEAGRTSCPCGGGLATKHGATSFQDCETRVCQSPGHFNTTTHRCIRCPVGTQPEFG 780  
 Db 719 PEAGRTSCPCGGGLATKHGATSFQDCETRVCQSPGHFNTTTHRCIRCPVGTQPEFG 778  
 Qy 781 KNNVCSCPGNTTDFDGSNTNITCKNRRCGGELGDTGYIESPNYGNYPANTECTWTIN 840  
 Db 779 KNNVCSCPGNTTDFDGSNTNITCKNRRCGGELGDTGYIESPNYGNYPANTECTWTIN 838  
 Qy 841 PPPKRRILIVVPEIFLPIEDDCDGLVMKRTSSNSVTYETCQTYERPIAFTSRSKLW 900  
 Db 839 PPPKRRILIVVPEIFLPIEDDCDGLVMKRTSSNSVTYETCQTYERPIAFTSRSKLW 898  
 Qy 901 IQPKSNEGSARGQVPPVYVYDQYQELIEDIVRDGRLYASENHQELKOKLIKALFDV 960  
 Db 899 IQPKSNEGSARGQVPPVYVYDQYQELIEDIVRDGRLYASENHQELKOKLIKALFDV 958

QY 961 LAHPQNYFYKTAQESREMPRSPFIRLLRSKVSRLPYK 999  
 DB 959 LAHPQNYFYKTAQESREMPRSPFIRLLRSKVSRLPYK 997

RESULT 8

ID ABP56757 standard; Protein; 964 AA.  
 XX ABP56757;

DT 31-MAR-2003 (first entry)

DE Human CEGP1 protein SEQ ID NO:5.

KW Human; secreted protein; epidermal growth factor protein subfamily;  
 KW epidermal growth factor; EGF; gene therapy; therapeutic; drug screening;  
 KW biological activity; immune response.

OS Homo sapiens.

XX NC02002101080-A2.

XX 19-DEC-2002.

XX 07-MAY-2002; 2002MO-US22278.

XX 16-MAY-2001; 2001US-0855824.

XX (PEKE ) PE CORP NY.

PA (GONG/) GONG F.

PA (DFRA/) DI FRANCESCO V.

PI Gong F, Di Francesco V, Yan C, Beasley EM;

XX WPI: 2003-167407/16.

XX New isolated human secreted peptides, useful for diagnosing or treating  
 PT a disease characterized by an absence of, inappropriate or unwanted  
 PT expression of the secreted protein, and in drug screening assays -

XX Disclosure; Fig 2E-F; 185pp; English.

XX The present sequence encodes a human secreted peptide (I) related to  
 CC the epidermal growth factor (EGF) protein subfamily. (I) can be used  
 CC in gene therapy. (I) and the nucleic acid molecules encoding (I) can  
 CC be used as models for the development of human therapeutic targets, to  
 CC aid in the identification of therapeutic proteins, and to serve as  
 CC targets for the development of human therapeutic agents. (I) may be  
 CC used in drug screening assays, in assays to determine the biological  
 CC activity of the protein, to raise antibodies or to elicit another immune  
 CC response, as a reagent in assays designed to quantitatively determine  
 CC levels of the protein in biological fluids, or as markers for tissues  
 CC in which the corresponding protein is preferentially expressed. (I) can  
 CC also be used for diagnosing or treating a disease or disorder  
 CC characterised by an absence of, inappropriate or unwanted expression of  
 CC the protein. (I) is located to human chromosome 22. The present sequence  
 CC represents human CEGP1 protein which is given in comparison with (I)  
 CC in the exemplification of the present invention.

XX Sequence 964 AA;

Query Match 96.8%; Score 5428; DB 24; Length 964;  
 Best Local Similarity 100.0%; Pred. NO. 0;  
 Matches 964; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 RGRAAGPQEDVDECAQGLDDCHADALCONPTSYKSCXGVOGEGRQCEDIDECGNELN 95  
 DB 1 RGRAAGPQEDVDECAQGLDDCHADALCONPTSYKSCXGVOGEGRQCEDIDECGNELN 60  
 QY 96 GGCYHDCNLIPGNRYCTCFDGFMLAHGHNCLDVECLNNGGCOHTCVNVMGSYECCCK 155

DB 61 GGCYHDCNLIPGNRYCTCFDGFMLAHGHNCLDVECLNNGGCOHTCVNVMGSYECCCK 120  
 QY 156 EGFFLSDNQHTCTHRSEGLSCVNKHGCSHIKKEAPRGSVACECPGFELAKNQDRCIL 215  
 DB 121 EGFFLSDNQHTCTHRSEGLSCVNKHGCSHIKKEAPRGSVACECPGFELAKNQDRCIL 180  
 QY 216 TCNHNNGSCGSCDCTADGPECSCCHPOYKMHDTGRSCLEREDTVLEVTESNTTSVVDGDK 275  
 DB 181 TCNHNNGSCGSCDCTADGPECSCCHPOYKMHDTGRSCLEREDTVLEVTESNTTSVVDGDK 240  
 QY 276 RYKRLLMETCAVNGGCDRTCKDITSTGVHSCSPVGFITLQDGTCKDIDECOTRNGGCD 335  
 DB 241 RYKRLLMETCAVNGGCDRTCKDITSTGVHSCSPVGFITLQDGTCKDIDECOTRNGGCD 300  
 QY 336 HFCXNIVGSFDCGCKGFKJLTDEKSCQVDDECSLORTQHSNCINHPGTPACACNRYTL 395  
 DB 301 HFCXNIVGSFDCGCKGFKJLTDEKSCQVDDECSLORTQHSNCINHPGTPACACNRYTL 360  
 QY 396 YGFTHCGDTHRECSINNGGCOQVNTVGSVECOCHPOYKJHWNKDCVEYKGLPTSVSP 455  
 DB 361 YGFTHCGDTHRECSINNGGCOQVNTVGSVECOCHPOYKJHWNKDCVEYKGLPTSVSP 420  
 QY 456 RVSUHCCKSGGCGCFLRCHSG:HLSSDVTITRTSVTFKLNKCKSLKNAELFPEGLRPA 515  
 DB 421 RVSUHCCKSGGCGCFLRCHSG:HLSSDVTITRTSVTFKLNKCKSLKNAELFPEGLRPA 480  
 QY 516 LPEKSSSVKESFRYVNLTCSSGKQVPCAPGRPTPKEMFITVEFELETNKEVTASCLLS 575  
 DB 481 LPEKSSSVKESFRYVNLTCSSGKQVPCAPGRPTPKEMFITVEFELETNKEVTASCLLS 540  
 QY 576 CIVKRTKRLRKA:RTLRKAVHRECEFLQ:SGMNLVDVAKKPPRTERCAESCGVGQGHAE 635  
 DB 541 CIVKRTKRLRKA:RTLRKAVHRECEFLQ:SGMNLVDVAKKPPRTERCAESCGVGQGHAE 600  
 QY 636 NQVSCRAGTYVDGARERC:LCPNGT:FNNEEGOMTCPECPFGNSGALKTPPEANMSFG 695  
 DB 601 NQVSCRAGTYVDGARERC:LCPNGT:FNNEEGOMTCPECPFGNSGALKTPPEANMSFG 660  
 QY 696 GLCOFGYSADGAFQCALGTFOEAGRTSCFPGGGGLATKHGATSPDCETRVQCS 755  
 DB 661 GLCOFGYSADGAFQCALGTFOEAGRTSCFPGGGGLATKHGATSPDCETRVQCS 720  
 QY 756 PGHFNITTHRCIRCPVGTYPBFGKNCVSCFGNTTDFDGSNTITQCKNRRCGGELGD 815  
 DB 721 PGHFNITTHRCIRCPVGTYPBFGKNCVSCFGNTTDFDGSNTITQCKNRRCGGELGD 780  
 QY 816 FTGVIESPNYPGNYPANTECTWTINPPPKRILIVVPEIFLPEDDCGDLVNRKTSNN 875  
 DB 781 FTGVIESPNYPGNYPANTECTWTINPPPKRILIVVPEIFLPEDDCGDLVNRKTSNN 840  
 QY 876 SVTTYETCCYVERPIATSFESKMLWQFKSNEGSARGPOVYVYDDEYOEELIETVRD 935  
 DB 841 SVTTYETCCYVERPIATSFESKMLWQFKSNEGSARGPOVYVYDDEYOEELIETVRD 900  
 QY 936 GRLYASENHQEI:KDKKL:KALFDVLAHPQNYKYTAQESREMPRSPFIRLLRSKVSRL 995  
 DB 901 GRLYASENHQEI:KDKKL:KALFDVLAHPQNYKYTAQESREMPRSPFIRLLRSKVSRL 960  
 QY 996 RPYK 999  
 DB 961 RPYK 964

RESULT 3  
 AAC16642  
 ID AAC16642 standard; Protein; 919 AA.  
 XX  
 AC AAC16642;  
 XX  
 DT 10-MAY-2003 (first entry);  
 XX Human extracellular messenger protein #2.  
 XX



```

PD 18-JUL-2002.
XX
XX 21-DEC-2001; 2001WO-US0548.
XX
XX 21-DEC-2000; 2000US-0747371.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Gish KC, Mack D;
XX
XX WPI; 2002-566749/60.
XX
XX
XX Diagnosing breast cancer comprises comparing the gene expression of a
XX breast cancer modulating protein in a test breast tissue and unaffected
XX breast tissue of the same or a different patient -
XX
XX
XX Disclosure; Fig 4A-B; 60pp; English.
XX
XX The invention relates to diagnosing breast cancer that involves
XX determining the expression of a gene encoding breast cancer modulating
XX protein (BCO2) or a fragment of it in a breast tissue of an individual,
XX and comparing the expression of the genes from a second normal breast
XX tissue from the individual or a second unaffected individual, where a
XX difference in expression indicates the individual has breast cancer. The
XX method is used for diagnosing breast cancer. Other new methods are
XX provided for screening drug candidates, screening for bioactive agents,
XX and evaluating the effect of a candidate breast cancer drug. An antibody
XX to BCO2 can be used to inhibit the bioactivity of BCO2 or neutralise it's
XX effect and can be used to treat breast cancer. The antibody can also be
XX used to localise a therapeutic group to breast cancer tissue, which can
XX be used to treat breast cancer. Antisense molecules are used to inhibit
XX breast cancer in a cell. BCO2 protein and polynucleotides are used to
XX elicit an immune response in an individual. BCO2 is also used to
XX determine the prognosis of an individual with breast cancer. The
XX present sequence represents the mouse BCO2 orthologue protein.
XX
XX Sequence 997 AA;
XX
XX
XX Query Match 90.2%; Score 5056; DB 23; Length 997;
XX Best Local Similarity 89.2%; Pred. No. 8e-306;
XX Matches 891; Conservative 47; Mismatches 59; Indels 2; Gaps 1
XX
XX
XX 1 MGVAGNRPRGAWAVLLLLLLLLPPLLLLLAGAVPPGRGAAGPOEDVDCAQGLDDCHADA 60
XX |||||
XX 1 MGVAGCGRPREARALLLLLLLLLLPP--LLAAAVPPDRGLTNGPSEDVDECAQGLDDCHADA 58
XX
XX 61 LCQNTPTSYKCSCKPGYGBGRQCEDIDECNELNGCGVHDCLINFGNVRCCFGDGMFLA 120
XX
XX 59 LCQNTPTSYKCSCKPGYGBGRQCEDMDCONTUNGCGVHDCLINFGNVRCTCFDGMFLA 118
XX
XX 121 HDGHNCLDVECLNNGGCQHTCVNVMGSECCCKGFFLLSDQHTCIHRSEEGLSCKMK 180
XX
XX 119 HDGHNCLDMECLNNGGCQHCITNVIGSEYCKCKGFFLLSDNQHTCIHRSEEGLSCKMK 178
XX
XX 181 DHGCSHI CKEAPRGSVACECPGFELAKNQDRDCLITCNHNGGCGQHSDDTDAGPECSCH 240
XX
XX 179 DHGCGHI CKEAPRGSVACECPGFELAKNQKDCILITCNHNGGCGQHSCEDTAEGPECSCH 238
XX
XX 241 POKYKHTDGRSLEREDTVLEVTSENSTSVVDGPKRYKRLMETCAVNGGCDRTCKDT 300
XX
XX 239 PYYRLHADGRSLCEGTGLEBTSNATSVADGDKRVARRLLMETCAVNGGCDRTCKDT 299
XX
XX 301 STGVHSCSPVGTLLQDGTKDKDIDECOTRNGGCDHFCKNIVSFGCGCKKGFKLITDEK 360
XX
XX 299 STGVHSCPTGFTLLQVDGKTKDKIDECOTRNGGCGNHFCKNIVSFGDCSKKGFKLITDEK 358
XX
XX 361 SCQDVEDCSLDRTCDHSCINHPGTFCACNRGYTYLGFTHCGDTNBSGINNGCQGVN 420
XX
XX 359 SCQDVEDCSLENTCDHSCINHPGTFCACNPGTYLSTFTHCGDTNECSVNGSCQGVN 418
XX
XX 421 TVGSEYECQHPGYKLUHWKXKUCVEVKGLLPTSVSPRVSJLHCGKSGGGDGCFLRCHSGIHL 480
XX
XX 419 TVGSEYECQHPGYKLUHWKXKUCVEVKGPPTSMTPRVSJLHCGKSGGGDRJLRCRSGIHL 478

```

[illegible]

RESULT 11	
ABP56758	
ID	ABP56758 standard; Protein; 957 AA.
XX	
AC	ABP56758;
XX	
DT	31-MAR-2003 (first entry;
XX	
DE	Mouse CEGP1 protein SEQ ID NO:6.
XX	
XX	Human; secreted protein; epidermal growth factor protein subfamily;
KW	epiderma. growth factor; EGF; Gene therapy; therapeutic; drug screening;
KW	biological activity; immune response.
XX	
CS	Mus musculus.
XX	
PN	W0200210:03C-A2.
XX	
PC	19-DEC-2002.
XX	
PF	07-MAY-2002; 2002WC-US22278.
XX	
PR	16-MAY-2001; 2001US-C855824.
XX	
PA	PEKE ; PE CORP NY.
PA	GONG/; GONG F.
PA	CFRA/; DI FRANCESCO V.
XX	
XX	Gong F, Di Francesco V, Yan C, Beasley BM;
PI	
XX	WPI; 2003-167407:16.
XX	
XX	
PT	New isolated human secreted peptides, useful for diagnosing or treating
PT	a disease characterized by an absence of, inappropriate or unwanted
PT	expression of the secreted protein, and in drug screening assays -

PS	XX	Disclosure: Fig 2f-G; 185pp; English.	
XX	XX	The present sequence encodes a human secreted peptide (I) related to	
XX	CC	the epidermal growth factor (EGF) protein subfamily. (I) can be used	
CC	CC	in gene therapy. (I) and the nucleic acid molecules encoding (I) can	
CC	CC	be used as models for the development of human therapeutic targets, to	
CC	CC	aid in the identification of therapeutic proteins, and to serve as	
CC	CC	targets for the development of human therapeutic agents. (I) may be	
CC	CC	used in drug screening assays, in assays to determine the biological	
CC	CC	activity of the protein, to raise antibodies or to elicit another immune	
CC	CC	response, as a reagent in assays designed to quantitatively determine	
CC	CC	levels of the protein in biological fluids, or as markers for tissues	
CC	CC	in which the corresponding protein is preferentially expressed. (I) can	
CC	CC	also be used for diagnosing or treating a disease or disorder	
CC	CC	characterised by an absence of, inappropriate or unwanted expression of	
CC	CC	the protein. (I) is located to human chromosome 22. The present sequence	
CC	CC	represents mouse CEGP1 protein which is given in comparison with (I)	
CC	CC	in the exemplification of the present invention.	
XX	SQ	Sequence 957 AA;	
		Query Match 88.0%; Score 4935; DB 24; Length 957;	
		Best Local Similarity 90.2%; Pred No. 2.5e-298;	
		Matches 862; Conservative 46; Mismatches 48; Indels 0; Gaps 0;	
QY	44	EDVDECAQGLDCHADALCQNTFTSYKSCKPGYQSGRQCEDEDECQNELNGCVHDC	103
Db	2	EDVDECAQGLDCHADALCQNTFTSYKSCKPGYQSGRQCEDEDECQNELNGCVHDC	61
QY	104	NIPGNRYCTCFGFMHLADHGNCLDVECLNNGCCQHTCVNMGYSYCCCKEFLSDN	163
Db	62	NIPGNRYCTCFGFMHLADHGNCLDVECLNNGCCQHTCVNMGYSYCCCKEFLSDN	121
QY	164	QHTCHRSSEGLSCNKHGCHICKAPRGSVACCEPGFELAKNQCOTLTCHNGGG	223
Db	122	QHTCHRSSEGLSCNKHGCHICKAPRGSVACCEPGFELAKNQCOTLTCHNGGG	181
QY	224	QHSQDDTADGPECSCHPOYKMTDORSCLEREDTVLEVTESNTTSVVDGKVRKRLLM	283
Db	182	QHSQDDTADGPECSCHPOYKMTDORSCLEREDTVLEVTESNTTSVVDGKVRKRLLM	241
QY	284	ETCAVANGGCDRTCKTSTGVHSCSPTFTLQVCGTKCDIDECQTRNGGCHFHCKNTVG	343
Db	242	ETCAVANGGCDRTCKTSTGVHSCSPTFTLQVCGTKCDIDECQTRNGGCHFHCKNTVG	301
QY	344	SFDCGCKGFKLLTDEKSCQDVDECSLDRTCQHSCHNHPGTFACACNRYGLYGFTHCGD	403
Db	302	SFDCGCKGFKLLTDEKSCQDVDECSLDRTCQHSCHNHPGTFACACNRYGLYGFTHCGD	361
QY	404	TNECSINNGGCGQVNTVGSYECQCHPGYKLFHWNKDCVVKGLLPTSVSPRSLHCGK	463
Db	362	TNECSINNGGCGQVNTVGSYECQCHPGYKLFHWNKDCVVKGLLPTSVSPRSLHCGK	421
QY	464	SGGDCGFLRCHSGHLSDVNTTIRTSVTFKLENGKSLKNAELFEGRLPALPEHSSV	523
Db	422	SGGDCGFLRCHSGHLSDVNTTIRTSVTFKLENGKSLKNAELFEGRLPALPEHSSV	481
QY	524	KESFYVNLTCSSGKQVPCAPGRSPTEKMFITVEFELTNCKEVTASCDLSICVKTREK	583
Db	482	KESFYVNLTCSSGKQVPCAPGRSPTEKMFITVEFELTNCKEVTASCDLSICVKTREK	541
QY	584	RURKARTLRKAVHREQFLQLSGMNLDAKPPPTSERQASCGVGQGHAEVQVSCRA	643
Db	542	RURKARTLRKAVHREQFLQLSGMNLDAKPPPTSERQASCGVGQGHAEVQVSCRA	601
QY	644	GTYDGAERCLICPNTQNEGQWTCPCPRPGNSGALKTPPEANNSECCGLQCPGEY	703
Db	602	GTYDGAERCLICPNTQNEGQWTCPCPRPGNSGALKTPPEANNSECCGLQCPGEY	661
QY	704	SADGFAPCOLCALGTFTQPEAGRTSCFPCCGGGLATHQGATSFQDCETRVCQSPGHFYNTT	763
Db	662	SADGFAPCOLCALGTFTQPEAGRTSCFPCCGGGLATHQGATSFQDCETRVCQSPGHFYNTT	721

```

Query Match.          65.6%; Score 3675.5; DB 24; Length 988;
Best Local Similarity 62.4%; Pred. No. 4.7e-250;
Matches 654; Conservative 113; Mismatches 152; Indels 119; Gaps 14;

10 GAA---WAVLLILLPULLLAGAVPGRGRAAG-----PQEDVBCAGLDDCHADAL 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2  GAAAVRHLVCVLLAL-----GTRGLAGGSLPGSVVDVDECSEGTDDCH1DAI 49

62 QNTPTS YKCSKPVGQEGROCEDI DCGNE -LNGCGVFDCLNI PGNVACTCPGPFMLA 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
50 QNTPKSYKLCRPGYKGGKOCEDI DCENDY YNGCGVHECNI PGNVRC-CFPGFMLA 109
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

121 HDGNCLLDVDECENNGGQCHTCVNMVGSYECCKEFFFSDNHQTCIHRSEBGLSCNMK 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
110 HDGNCLLDVDECODNNGGQQICVNMVGSYECQCHSGFFLSDNHQTCIHRSEBGLSCNMK 169
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

181 DHCSSHICKEAPRGSVACERPGFELANQBDCLTCNHNGGCGQHSCEDTADGPECSCH 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
170 DHCAGHICRETPKGGVACDRPGFLAQKDCLTCTNYNGGCGHSCEDTDTGTCTCSH 229
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

241 PQYKMHITDGRSLEREDTVLEVTESNTTSVVDGDKVRRLMETCAVNGGCDRTCKDT 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
230 QKYALHSDGRTCI-----ETCAVNGGCDRTCKDT 259

301 STGVHSCSPVGFLLDGLKTKDIDECOTRNGGCHFCKNIVSGFCGCKKFKKLLTDEK 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
260 ATGVRCSCPVGFTLPBGKTKDINECLVNVNGGCHFCRNTVSGFCGCKRKGKLLTDER 319
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

361 SCQDVDECLDRDCHSCINHPGTFCACNRGYTLGYTHCGTNECSINNGCQGVGVN 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
320 TCQDIDECSPERTCDHICINSPGSFQCLCHRGYLYGTHCGDVDECSXNGSCDQGVN 379
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

421 TVGSYECQCHPGYKJHWNKDCVEV-KGLLPTSVSPVSV-HCGKSGGDCGFLRC----- 474
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
380 TKGSYECVCPGPRRLHWNGKCVETGKCLSAKTSAPRAQCSKAGVSECSFUSPAUHL 439
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

475 -----HSGIHL-----SDVTIRTSVTFKLNEGKSL 502
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
440 FVPDSSENSVYLCGVPGPQKALQKRNGTSGLGPGSCSDAPTPIKQKARFKIRAKCHL 499
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

503 KNAELPEGLPALPEKHSVSESR-----YVNLTCSSGKVQVCAQRSDTRKE 552
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
500 R-----PHSQAKETARQPLDHCHTVFTLKCDSSKK--RRGRKSPKSE 544
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

553 M-FITVEFELETNQEVTASCDLSGIVKRTBKRKAIRTLKAVAHREOFHLQSGMND 611
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
545 VSHITAEFIETKWEASDTCEADCLRKPRABSOAAIKTLRKSIGRQFYVQVSGTEYE 604
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

612 VAKKPPRTSRQAESCGVGQHAENQVSCRAGTYDGAERECILCPNGTFQNEGQVTC 671
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
605 VAQRPKALBGGQ-ACGAGVQLDQSKVACGPGTHFGELGQCVSWPGTYQDMEGQLSC 663
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

672 EPCRPNGSALKATPBAWNVKSGGLGQPGYVSADGAPQOLCALGTFQPEAGRTSCPCC 731
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
664 TPCP-----SSDGLGLPGARVYSCGQSPGFSFADGKPCQACPVSTYQPEQRTGCPFC 720
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

732 GGGLATKHQCATSFQCECTRVOCSPGHFVNTTHRC-RCPVGTYPQEPFGNQCVCSPGNT 791
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
721 GGGLLTKHEGTSFQDCEAKVCSFGHHYNTTHRCIRCPVGTYPQEPFGNQCVCSPGNT 780
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

792 TTDPDGSNTITQCNRRCGELGDFTVYESPNYPGNVPAWNTCTMTINPPKRRLLVY 851
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
781 STDPDGSNTVTHCNRQHCGLGQVTVIESPNYPGVYPAWNTCTMTINPPKRRLLVY 840
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

852 PEIPLIEDCGGVLYNRKTSNNSVTVETCTQVERPIAFTSRKSLMTQFKNEGNSA 911
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
841 PEIPLIEDCGGVLYNRKTSNNSVTVETCTQVERPIAFTSRKSLMTQFKNEGNSG 900
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

912 RGFQVPVTVYDEYQELIEDIVDRGLYASENHOEILKDKKLKALFDVLIAHPQNYFY 971
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
901 KGFQVPVTVYDEYQELIEDIVDRGLYASENHOEILKDKKLKALFDVLIAHPQNYFY 960
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Cy 972 AQSREMFPSRIRLLRSKVSRFLRPYK 999  
[|||||] [|||||] [|||||] [|||||]  
Db 96: AQSEKMFPSRIKLRLSKVSRFLRPYK 988

RESULT J3  
ABC04645  
ID ABC04649 standard; Protein: 988 AA.  
XX AC ABJC4648;  
XX XX  
DT 11-OCT-2002 (first entry)  
XX XX  
DE Protein of NOVX ab SEQ ID No 16.  
XX XX  
KW Cytostatic; antidiabetic; anorectic; metabolic; neuropathic; neurotropic; antilipaeamic;  
KW neuroprotective; anti-parkinsonian; anticorvulsant; cerebroprotective; cerebropsective;  
KW tranquilliser; neuroleptic; antidiabetic; antitumor; antitumor; antineoplastic;  
KW anti-Hiv; anti-allergic; antirheumatic; antithrombotic; NOXV; diabetes;  
KW metabolic disorder; obesity; infectious disease; Alzheimer's disease;  
KW anorexia; neurodegenerative disorder; Parkinson's disorder; obesity;  
KW immune disorder; haematopoietic disorder; dyslipidaemia; chronic disease  
KW metabolic syndrome X; wasting disorder; cancer; neurological disorders; gout;  
KW epilepsy; stroke; mental disorder; schizophrenic disorders; goiter;  
KW vesicular transport; cystic fibrosis; gastrointestinal disorder;  
KW diabetes mellitus; ulcerative colitis; AIDS; allergic reaction;  
KW multiple sclerosis; rheumatoid arthritis; transgenic animal;  
Gene therapy.  
XX KW  
XX Unidentified.  
OS OS  
XX XX  
PN WQ20C2464C9-A2.  
XX XX  
PD 13-JUN-2002.  
XX XX  
PF 06-DEC-2001; 2001WO-US46586.  
XX XX  
PR 06-DEC-2000; 200CUS-231660P.  
PR 12-DEC-2000; 200CUS-355029P.  
PR 08-JAN-2001; 200LUS-260326P.  
PR 24-JAN-2001; 200LUS-263800P.  
PR 28-FEB-2001; 200LUS-269942P.  
PR 24-APR-2001; 200LUS-286193P.  
PR 20-AUG-2001; 200LUS-313627P.  
PR 12-SEP-2001; 200LUS-318712P.  
XX XX  
PA (CURA-) CURAGEN CORP.  
XX XX  
PI Guo X, Li L, Patturajan M, Shimkets RA, Casman SJ, Malyankar UM,  
PI Chernev VT, Verret CAM, Spytek KA, Shenoy SG, Alschrook JP;  
PI Edinger S, Peyman SA, Stone DC, Ellettan K, Gargoli EA;  
PI Boldog FJ, Colman JD, Eisen AJ, Liu X, Padigaru M. Spaderna SK;  
PI Zerhusen BD;  
XX XX  
WP1: 2002-547774/58.  
DR N-PSDB; ABT05461.  
XX XX  
PT Novel isolated polypeptide, designated NOVX, useful for treating or  
PT preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and  
PT metabolic, neurodegenerative, immune and hematopoietic disorders -  
XX XX  
CS Claim 1; Page 69; 42pp; English.

The invention relates to an isolated polypeptide, designated NOVX, comprising a sequence fully defined in the specification. The isolated protein, its encoding polynucleotide or an antibody created from the protein is useful in the manufacture of a medicament for treating a syndrome associated with a human disease, preferably a NOVX-associated disorder, or for treating or preventing a NOVX-associated disorder in a subject, preferably human. The isolated protein, its encoding polynucleotide or an antibody created from the protein are also useful for treating or preventing metabolic disorders, diabetes, obesity, infectious disease, anorexia, neurodegenerative disorder, Alzheimer's

CC disease, Parkinson's disorder, immune disorders, haematopoietic  
 CC disorders, and various dyslipidaemias, metabolic disturbances associated  
 CC with obesity, the metabolic syndrome X, wasting disorders associated with  
 CC chronic diseases, and cancer. The isolated protein, its encoding  
 CC polynucleotide or an antibody created from the protein are useful for  
 CC treating or preventing neurological disorders such as epilepsy, stroke,  
 CC mental disorders including mood, anxiety, schizophrenic disorders,  
 CC disorders of vesicular transport such as cystic fibrosis, diabetes  
 CC mellitus, goiter, gastrointestinal disorders including ulcerative  
 CC colitis, other conditions associated with abnormal vesicle trafficking  
 CC including AIDS, allergic reactions, multiple sclerosis and rheumatoid  
 CC arthritis. A cell comprising the vector of the invention is useful for  
 CC producing non-human transgenic animals. The polynucleotide of the  
 CC invention can be used to treat disorders by gene therapy. This sequence  
 CC represents one of the isolated NOVX proteins of the invention.  
 XX  
 SQ Sequence 988 AA;

Query Match 65.3%; Score 3662.5; DB 23; Length 988;  
 Best Local Similarity 62.2%; Pred. No. 3e-219;  
 Matches 652; Conservative 113; Mismatches 164; Indels 119; Gaps 14;

QY 10 GAA---WAVLLELLPALLLAGVPPGCRAG-----PQEDVDECAOGLDCHADAL 61  
 DB 2 GAAYRHLVCLLAL-----GTRGLAGSGLPQGVVDVDESGTDOCHDAI 49

QY 62 QNTPTSYKCSKPGYQEGRQCEIDECGNE-LNGGCVHDCILIPGNYRCTCFDGFMLA 120  
 DB 110 HGHNCLDVECDQNNNGCQQTIVNMGSGYECQCHSGFFSDNQHTCIHRSNMGNCMK 169

QY 121 HGHNCLDVECDQNNNGCQQTIVNMGSGYECQCHSGFFSDNQHTCIHRSNMGNCMK 180  
 DB 110 HGHNCLDVECDQNNNGCQQTIVNMGSGYECQCHSGFFSDNQHTCIHRSNMGNCMK 169

QY 181 DHGCSHICKEAPRGVACECRGPELAKNQDCILTNHNGGCGHSCDITADGPECSCH 240  
 DB 170 DHGCAHICRETPKGVACDCRGPPLAQNQDCILTNHNGGCGHSCDITADGPECSCH 229

QY 241 POKMHTGGRSCLEREDTIVLETSNTTSVVDGDKVRKRLMETCAVNGGCDRCCKT 300  
 DB 230 QKYAPHSDGRTCI-----ETCAVNGGCDRCCKT 259

QY 301 STGVHCSPVGTLLDGLDKTKIDECOTRNGGDHFCNTVGSFDCGCKGKGLLTDEK 360  
 DB 260 ATGVRCSPVGTLLDGLDKTKIDECOTRNGGDHFCNTVGSFDCGCKGKGLLTDEK 319

QY 361 SCQVDECSLORTDHSINCINHPGTACACNRGYLYGTHCGTNECSINNGGCGQCVVN 420  
 DB 320 TCQDIDEGSFERTCDHICINSPGSPQCLCHRGYLYGTHCGVDVDECSNMGSDQCVN 379

QY 421 TVGSYECOCCHPGYKLHNKKOCVEV-KGLLPSTVSFVSLHCKSGGDDGCFRLC----- 474  
 DB 380 TKGSEYECVCPGRRUHRNKKOCVERGKCLSAKTSIPRAQUSCKAGVSECFLSCPAHTL 439

QY 475 -----HSGIHLS-----SDVTIRTSVFKLNEKCSL 502  
 DB 440 FVPDSENSYVLSCGVPGQKALQKRNCTSSGLSPSCSDAPTTPKCKAREKINDAKHL 439

QY 503 KNABLFPEGLRALPEKUSVKEFR-----YVNLTCSSGKQVPGAPRSTPKPE 552  
 DB 500 R-----PHSOARAKETARQPLLDHCVTFVTLKCDSSKK--RRGRKSPSKE 544

QY 553 M-FITVEFELETNQEVATSCDLSICIVARTEKRLKRAKRT-RKAVHREQFHLQSLGNJD 611  
 DB 545 VSHI-FAEFIEI-WMEASDTCBADCLRKRAQSLQAALTKLXKIGRQYVQVUSGTEYE 604

QY 612 VAKKPPRTSERAECSGGVGHAEQVSCAGTYVDGARERCILCPNGTTFQNEGGMTC 671  
 DB 605 VAQRPALALEGGQ-ACGAGQVLQDSKCVACPGTHFGGELGQCVPCMPFTYQWMEGLSC 663

QY 672 EPCPRFGNSGALKTPAEMNMSCGGLGCPGEYSADGAPCOLCALGTQFQEAGRSTCFPC 731  
 DB 664 TPCP-----SSDGLGLPGARNVSECGGQCGPGFFSADGFKPCACPVGTQVPEPGRTCFPC 720

QY 732 GGLIATKHGATSFQDCETRVQSPGHFYNTTTHRCIRCPVGTYQPEFGKNCVSCPNT 731  
 DB 721 GGLLTTHRGHTSFQDCETRVQSPGHFYNTTTHRCIRCPVGTYQPEFGKNCVSCPNT 780

QY 792 TTDGSGTNITQCKNVRPCGSGELGDTGY-ESPNYGVNYPAN-ECTWTINPPKRRILIV 851  
 DB 781 STDGSGTNITQCKNVRPCGSGELGDTGY-ESPNYGVNYPAN-ECTWTINPPKRRILIV 840

QY 952 PEIPLPIEDCGGVLYNWKTSNNSVTTTCTYERPIAF-SRSKKKMIQPKSNEGNSA 911  
 DB 941 PEIPLPIEDCGGVLYNWKTSNNSVTTTCTYERPIAF-SRSKKKMIQPKSNEGNSA 900

QY 912 RGFQVPVVTYDQYQELIEDIVDGRLYASENHCEILKDKKLILKAFDVLAKPQNYFKYT 971  
 DB 901 KGFQVPVVTYDQYQELIEDIVDGRLYASENHCEILKDKKLILKAFDVLAKPQNYFKYT 960

QY 972 AQESREMPFSPFILLASKSVSRFLRPYK 999  
 DB 961 AQESREMPFSPFILLASKSVSRFLRPYK 988

RESULT 14  
 AAU79172 standard; Protein; 1006 AA.  
 XX  
 AC AAU79172;  
 CT 02-JUL-2002 (first entry)  
 XX  
 DE Human MEGF/FIBRILLIN-like protein NOV8.  
 XX  
 KW Human; MEGF/FIBRILLIN-like protein; receptor; NOVX; NOV8;  
 KW vaccine; cancer; tumour; bone disorder; avascular necrosis; allergy;  
 KW haematopoietic disorder; immune disorder; endometriosis; renal disease;  
 KW infection; inflammatory disease; lung disease; scleroderma; ataxia;  
 KW bowel disease; appendicitis; blood disorder; cardiovascular disorder;  
 KW graft versus host disease; GVHD; lymphoedema; brain disorder;  
 KW ocular disorder; hepatitis C virus infection; cardiac disorder;  
 KW autosomal dominant deafness; DFNA-2; chromosome 6.  
 XX  
 GS Homo sapiens.  
 XX  
 FN W0200214368-A2.  
 XX  
 PD 21-FEB-2002.  
 XX  
 PF 16-AUG-2001; 2001WC-US25624.  
 XX  
 PR 16-AUG-2001; 2000US-225692P.  
 PR 16-AUG-2001; 2000US-225693P.  
 PR 16-AUG-2001; 2000US-225837P.  
 PR 18-AUG-2000; 2000US-226236P.  
 PR 18-AUG-2000; 2000US-226353P.  
 PR 22-AUG-2000; 2000US-227085P.  
 PR 23-AUG-2000; 2000US-227395P.  
 PR 24-AUG-2000; 2000US-227492P.  
 PR 24-AUG-2000; 2000US-227600P.  
 PR 14-MAR-2001; 2001US-275952P.  
 XX  
 FA {CURA-} CURAGEN CORP.  
 XX  
 PI Zethusen BD, Padigaru M, Spyttek KA, Spaderra SK, Gangolli EA;  
 PI Bastelli Z, Burgess CE, Matumder K, Shinkens R, Mishra V;  
 PI Vernet CAM, Szekeres ES, Grosse WX, Alcobrook CP, Liu X;  
 PI Gerlach VS, Ellerman K, Smithson G, Peyman J, Stone D;  
 PI MacDougall J;  
 DR WPI; 2002-329571/36.  
 DR N-PSDB; ABK48392.  
 XX  
 PT Novel cytoplasmic, nuclear membrane bound and secreted NOVX  
 PT polypeptides, useful for treating cancers and tumours, bone disorders,



PT Paget's disease, haematopoietic disorders, spinal diseases and immune  
 XX disorders  
 PS  
 XX Claim 1; Page 77; 234pp; English.  
 CC The present invention relates to new isolated NOVX polypeptides named  
 CC NOV1-NOV9. The invention can be used for identifying an agent (a cellular  
 CC receptor or downstream effector) that binds to the polypeptide. The  
 CC molecules of the invention are useful for treating or preventing  
 CC NOVX-associated disorders in humans. The antibody of the invention is  
 CC useful for determining the presence or amount of NOVX in a sample, and  
 CC for treating a pathological state in a mammal. The method of the  
 CC invention is useful for determining the presence of an amount of NOVX in  
 CC a sample which is used as a marker for cancerous cell or tissue type.  
 CC The molecules of the invention are useful in the manufacture of a  
 CC medicament for treating or preventing cancer, tumour, bone disorders,  
 CC avascular necrosis, allergy, haematopoietic disorders, immune disorders,  
 CC endometriosis, renal diseases, infections, inflammatory diseases, lung  
 CC diseases, scleroderma, ataxia, bowel diseases, appendicitis, blood  
 CC disorders, cardiovascular disorders, graft versus host disease (GVHD),  
 CC lymphoedema, brain disorders, ocular disorders, hepatitis C virus  
 CC infection, cardiac disorders and autosomal dominant deafness (DFNA-2).  
 CC The present amino acid sequence represents the human MEGF/FABRILIN-like  
 CC protein NOV8 that is one of the NOVX proteins described in the invention.  
 CC This sequence is encoded by the human MEGF/FABRILIN-like protein NOV8  
 CC gene located on chromosome 6.

Query Match 61.4%; Score 3443.5; DB 23; Length 1006;  
 Best Local Similarity 59.6%; Pred. No. 1.2e-205;  
 Matches 601; Conservative 150; Mismatches 217; Indels 41; Gaps 9;  
 4 AGRNPPGAAWVLLLLPPLLLAGAVPPGRRAGPQEDVDEACQGLDDCHADALCO 63  
 26 AAQKGEGARGLGAVLAAGPQL---GAIPGGGRCMGPLPDVDECVSTNDCHDAICQ 92  
 64 NTPSYKSCKPGVGGGROCEDIDECGNELNGCVHDCNIPGNVRCCTCFDGFMAHOG 123  
 83 NTPSYKICSGYTGDKCKVDCEEREDNACGVHDCVNIIGNVRCITCYDGFMAHOG 142  
 124 HNCILDVDELENNGGQHTCVNNGSYECCKGFFLSDNQHTCIHRSEGLSCMNKHG 183  
 143 HNCILDVDECAENGCGCQSVNMGSYECHECREGFFLSNQHTCIORPEGMNCKNHG 202  
 184 CSHICKAPRSVACRCRPFELAKNORDCILTCNNGGCGCHSCDDTAGDESCCHPY 243  
 203 CAHICRTPTKGIACRCRPFELTKNQDKLTCTNNGGCGCHTCDTTEGPRGCGCH:KP 262  
 244 KMTDGRSCLEREDTVLEVTESTNTTSVVDGDKRVKRL-----LMETCAVNGGCDORTCK 299  
 263 VLHTDGTCTI-----GERRLEQH:PTQAVSNETCAVNGGCDCKSKCH 303  
 299 DTSTGVHSCDPVGTLLQ:DKTKDKIDECQTRNGGCDHFCNIVGSDGCKKGFKLITC 358  
 304 DAATGVHCTCPVGMFLQFDRKTKIDIDECRLNNGGCDHI CRNTVGSFPECCKGYKLL:N 363  
 359 EKSQDVDECSLDRCTDHSCHNHSPTACACNRYGTYLGFTHCGDTNECSINNGGCGQVC 418  
 364 ERNQQIDECSDFTCTDHCICNTPGSQCLCHRGYLLYGT:HCQDVDECSINRGGRFGC 423  
 419 VNTVGSVECCQHPGY-KLHNKNDKCVKGLLPTSVSPRVSLHCGKSGGCGGFLRCHSG 477  
 424 INTPGSYQCTCPAGQGRLLHNGKCTEP:KCGQSPGASKAVLSNCRSGKXDTCALTCPS 483  
 478 IHLSDVTTIRTSVTFKLNKCKSLKVAELFPEG:----RPALEKHSVSKESFRVNL 533  
 484 IPEAAVLISIKQASFKIKAKRCLAGSENFPGDVPVGVDPKPSLFGGAPCSECQVTFHLK 543  
 534 CSSGKQVPCAPGRSTPKMEITVEFELETNOKVETASCDLSICVKTETKRLKKA:PTLR 593  
 544 CDSRKGKGRRRARPSCSQQLFLP-----DTHGHPFPASCGLPCLQRMRERLKGSLKMLR 599

QY 594 KAVHBEQPHLQLSSXNLDVAKKPPRTSERQA---ESCGVGGCGHAENCCVSCRAGTYDGA 650  
 DB 600 KSNQDRFLRLDAGLYELAHKFGLVAGRAEMESCRFGQHRADSPSVSCFGTYHYGQ 659  
 QY 651 REPCLICPNCTFQNERGOMTCEPCPRPGNSGALKTPEANMSECGGLCPQGEVSAPGFAP 710  
 DB 660 TCCVCPCEATPBERGQUSCD:CPREKQGSPPKPKHGAGPKCGOR :RRKHESEDGFAP 718  
 QY 711 COLCALGTFOPEAGRTSCFPCGGGLATKMGATSFQDCETRVQCSGPHFYNTTHRCIRC 770  
 DB 719 CQCPRGTYQFEAGRLCPFGCGGLATKHEGATSFQDCD:KVQCSGPHYNTTHRCIRC 778  
 QY 771 PVGTQPEFQNNVCSPGNTTDFPGSINI:TOCKNRRCGGELGDTGVIESPNYGNYP 830  
 DB 779 AMGSCQDFPQNCRCRCPGNT:STDFPGSIVAGCKNRQCGGELGDTGVIESPNYGNYP 838  
 QY 831 ANTECTWT:NPPPRRL:TVVPIFLPIEDDQGVYVMBKTSNNSVTTVETCQTVVERI 890  
 DB 839 AGVECT:NNINPPPKRLIIVVPIE:PSDECGDLVYMKNSPSSITVETCQTVVERI 898  
 QY 891 APTSRKALMIQKSNKSGARGFQVPPYVYDYOELLIEDVDRGLVASENHQBLKD 953  
 DB 899 APTARSKALMIQKSNKSGARGFQVPPYVYDYOELLIEDVDRGLVASENHQBLKD 958  
 QY 951 KKLJKA:FDVLAHPQRYKVTAGSEMEFPRSRIRLLRSKVSRLRPYK 999  
 DB 959 KKLKAF:FEVLAHPQRYKVTAGSEMEFPRSRIRLLRSKVSRLRPYK 1006  
 RESULT 15  
 ABU12093  
 CD ABU12093 standard; Protein; 974 AA.  
 XX  
 AC ABU12093;  
 XX  
 DT 17-FEB-2003 (first entry)  
 XX  
 DE Novel human epidermal growth factor-like protein #3.  
 KW Gamma-aminobutyric acid receptor-like protein; depression; stroke;  
 KW GABA receptor-like protein; Parkinson's disease; Huntington's disease;  
 KW Tourette's syndrome; amyotrophic lateral sclerosis; head trauma;  
 KW Alzheimer's disease; alcoholism; vigilance; anxiety; muscle tension;  
 KW epileptogenic activity; memory; cardiomyopathy; cancer; angiogenesis;  
 KW arhythmogenic right ventricular dysplasia; renal disease; diabetes;  
 KW Epidermal growth factor like protein; leukaemia; lupus; anaemia; ulcer;  
 KW haematopoietic stem and progenitor cell like protein; cirrhosis;  
 KW sulfolipase-like protein; cholangitis; hepatitis; hyperthyroidism;  
 KW developmental disorder; Syntaxin-like protein; myxoid liposarcoma;  
 KW asthma; Lambert-Eaton myasthenic syndrome; acute myeloid leukaemia;  
 KW transgenic animal.  
 XX  
 CS Homo sapiens.  
 XX  
 PN US2002123612-A1.  
 XX  
 PD CS-SEP-2002.  
 XX  
 PF CS-JUL-2001; 2001US-0898570.  
 XX  
 PR 19-APR-2000; 2000US-198293P.  
 PR 20-APR-2000; 2000US-198645P.  
 PR 25-APR-2000; 2000US-199476P.  
 PR 26-APR-2000; 2000US-199880P.  
 PR 26-APR-2000; 2000US-200024P.  
 PR 26-APR-2000; 2000US-200025P.  
 PR 09-JUN-2000; 2000US-210809P.  
 PR 03-JUL-2000; 2000US-215851P.  
 PR 17-JUL-2000; 2000US-218591P.  
 PR 11-AUG-2000; 2000US-224612P.  
 XX  
 PA :GERLACH V L.  
 PA :ELLE/ ELDERMAN K.



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2003, 11:30:24 ; Search time 92 Seconds  
(without alignments)  
2040.161 Million cell updates/sec

Title: US-09-747-371-2

Perfect score: 5605  
Sequence: 1 MGVRNRRFGAAWAVLLLLL.....PRSFRLRLRSKVSFLRPRYK 999

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/ACTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NPK\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5605	100.0	999	9	US-09-747-371-2
2	5605	100.0	999	15	US-10-177-293-45
3	5581	99.6	997	15	US-10-176-847-50
4	5428	96.8	964	12	US-09-855-824-5
5	5056	90.2	997	9	US-09-747-371-3
6	4935	88.0	957	12	US-09-855-824-6
7	3675.5	65.6	988	12	US-09-855-824-2
8	3675.5	65.6	988	15	US-10-195-142-2
9	3417.5	61.0	974	20	US-09-898-570-14
10	3417.5	61.0	974	11	US-09-839-446-14
11	3397	60.6	993	12	US-10-239-663-36
12	3394	60.6	1009	10	US-09-898-570-16
13	3394	60.6	1009	11	US-09-839-446-16
14	3290.5	58.7	964	12	US-10-239-663-58
15	3290.5	58.7	965	12	US-10-221-097-51

16	3282.5	58.2	911	12	US-09-855-824-4
17	2931.5	52.3	845	10	US-09-898-570-12
18	2931.5	52.3	845	11	US-09-839-446-12
19	2909	51.9	897	12	US-10-239-663-35
20	2819	50.3	785	15	US-10-195-142-8
21	2439.5	43.5	735	10	US-09-898-570-10
22	2439.5	43.5	735	11	US-09-839-446-10
23	2162	38.6	392	10	US-09-898-570-39
24	2162	38.6	392	11	US-09-839-446-39
25	1680	33.5	534	15	US-10-195-142-10
26	1826	32.6	467	12	US-10-221-097-50
27	1615.5	28.8	418	10	US-09-886-429-2
28	1615.5	28.8	418	15	US-10-162-435-21
29	1239.5	22.1	333	15	US-10-195-142-4
30	1091	19.5	421	12	US-10-017-161-16C6
31	955.5	16.8	250	15	US-10-195-142-6
32	978.5	17.5	286	15	US-10-195-142-12
33	945	16.6	387	12	US-10-239-663-57
34	932.5	16.6	208	9	US-09-764-853-683
35	932.5	16.6	211	11	US-09-764-881-114
36	927.5	16.5	224	9	US-09-764-853-848
37	927.5	16.5	224	11	US-09-764-881-162
38	681	12.1	776	14	US-10-600-512-8
39	681	12.1	959	14	US-10-600-512-10
40	635	11.3	915	10	US-09-909-320-34
41	635	11.3	915	10	US-09-909-320-34
42	635	11.3	915	10	US-09-905-201A-34
43	635	11.3	915	10	US-09-902-853-34
44	635	11.3	915	10	US-09-907-824-34
45	635	11.3	915	10	US-09-907-824-34

ALIGNMENTS

RESULT :

US-09-747-371-2  
; Sequence 2, Application US/59747371  
; Patent No. US20020006616A1  
; GENERAL INFORMATION:  
; APPLICANT: Gist, Kurt  
; APPLICANT: Mack, David  
; TITLE OF INVENTION: NO. US20020006616A1 Methods of Diagnosing Breast Cancer, Comp  
; TITLE OF INVENTION: Screening for Breast Cancer Modulators  
; FILE REFERENCE: A-69029/DJB/JJD  
; CURRENT APPLICATION NUMBER: US/09/747,371  
; CURRENT FILING DATE: 2000-12-21  
; PRIOR APPLICATION NUMBER: PCT/ US/00/06952  
; PRIOR FILING DATE: 2000-03-15  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 2  
; LENGTH: 999  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-747-371-2

Query Match:					100.0%; Score 5605; DB 9; Length 999;
Best Local Similarity					100.0%; Pred No. 0;
Matches 999; Conservative					0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	MGVAGNRFGAAWAVLLLLL	1	PPILLLAGAVPPGRGAAAGCEVDCAQGLDCHADA	60
Db	1	MGVAGNRFGAAWAVLLLLL	1	PPILLLAGAVPPGRGAAAGCEVDCAQGLDCHADA	60
Qy	61	LCQNTPTSYKCSKPCYQGEGRCCEDIDECGNELNGCVVHDCNLNIPGNVRCCTCFDGFMLA	120		
Db	61	LCQNTPTSYKCSKPCYQGEGRCCEDIDECGNELNGCVVHDCNLNIPGNVRCCTCFDGFMLA	120		
Qy	121	HDGHNCLDVEDEENGGCGQHTCVNVMGSYECCKEGFELSNQHTCIHRSEGLSCMNK	180		
Db	121	HDGHNCLDVEDEENGGCGQHTCVNVMGSYECCKEGFELSNQHTCIHRSEGLSCMNK	180		

Qy	181	DHGC	SHI	CKE	A	P	R	G	S	V	A	C	E	R	P	G	F	E	L	A	K	Q	R	D	C	I	L	T	N	H	G	N	G	C	O	H	S	C	D	T	A	D	P	E	C	S	C	H	240							
Db	181	DHGC	SHI	CKE	A	P	R	G	S	V	A	C	E	R	P	G	F	E	L	A	K	Q	R	D	C	I	L	T	N	H	G	N	G	C	O	H	S	C	D	T	A	D	P	E	C	S	C	H	240							
Qy	241	PQY	X	H	T	D	G	R	S	C	L	E	R	E	D	T	V	L	E	S	N	T	T	S	V	D	G	K	R	V	K	R	-	L	M	E	T	C	A	V	N	G	G	O	R	-	C	K	T	300						
Db	241	PQY	X	H	T	D	G	R	S	C	L	E	R	E	D	T	V	L	E	S	N	T	T	S	V	D	G	K	R	V	K	R	L	L	M	E	T	C	A	V	N	G	G	O	R	-	C	K	T	300						
Qy	301	STG	V	H	C	S	C	P	V	G	T	L	D	G	K	T	C	K	D	I	E	C	O	T	R	N	G	G	D	H	F	C	K	-	V	G	S	D	C	G	K	K	F	L	J	D	E	K	360							
Db	301	STG	V	H	C	S	C	P	V	G	T	L	D	G	K	T	C	K	D	I	E	C	O	T	R	N	G	G	D	H	F	C	K	-	V	G	S	D	C	G	K	K	F	L	J	D	E	K	360							
Qy	361	SQD	V	B	E	C	S	L	D	R	T	C	D	H	S	C	I	N	P	G	T	F	A	C	A	C	N	R	G	T	L	Y	G	T	H	C	G	T	N	E	S	I	N	N	G	C	Q	V	C	N	420					
Db	361	SQD	V	B	E	C	S	L	D	R	T	C	D	H	S	C	I	N	P	G	T	F	A	C	A	C	N	R	G	T	L	Y	G	T	H	C	G	T	N	E	S	I	N	N	G	C	Q	V	C	N	420					
Qy	421	TVGS	E	C	O	C	H	P	G	Y	K	L	H	N	K	D	C	V	E	V	K	L	L	P	T	S	V	S	P	R	V	S	L	H	C	G	S	G	G	D	G	F	L	R	C	H	S	I	H	480						
Db	421	TVGS	E	C	O	C	H	P	G	Y	K	L	H	N	K	D	C	V	E	V	K	L	L	P	T	S	V	S	P	R	V	S	L	H	C	G	S	G	G	D	G	F	L	R	C	H	S	I	H	480						
Qy	481	SSD	V	T	T	I	R	T	S	V	T	F	L	K	N	E	G	K	C	L	K	N	A	E	L	F	P	E	G	L	R	P	A	L	P	E	K	S	S	V	K	E	S	P	R	Y	N	L	T	C	S	G	K	Q	V	540
Db	481	SSD	V	T	T	I	R	T	S	V	T	F	L	K	N	E	G	K	C	L	K	N	A	E	L	F	P	E	G	L	R	P	A	L	P	E	K	S	S	V	K	E	S	P	R	Y	N	L	T	C	S	G	K	Q	V	540
Qy	541	PC	A	G	R	P	S	T	P	K	E	M	F	I	T	V	E	F	E	L	T	N	O	K	E	V	A	S	D	L	S	C	I	V	K	T	E	K	L	R	K	A	I	R	K	A	V	H	E	Q	600					
Db	541	PC	A	G	R	P	S	T	P	K	E	M	F	I	T	V	E	F	E	L	T	N	O	K	E	V	A	S	D	L	S	C	I	V	K	T	E	K	L	R	K	A	I	R	K	A	V	H	E	Q	600					
Qy	601	FH	L	Q	S	G	M	L	D	V	A	K	P	P	R	T	S	R	Q	A	E	S	C	G	V	G	G	H	A	E	N	C	V	S	C	R	A	G	T	Y	D	G	A	R	E	S	C	I								

## RESULT 2

US-101-177-293-45  
 ; Sequence 45, Application US/101/77293  
 ; Publication No. US20030124128A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Iillie, James  
 ; APPLICANT: Giatt, Karen  
 ; APPLICANT: Zhao, Xumei  
 ; APPLICANT: Gannavarpu, Marjula  
 ; APPLICANT: Kanatkar, Subhanghi  
 ; APPLICANT: Mertens, Maureen  
 ; APPLICANT: Myer, Vic  
 ; APPLICANT: Wang, Youzhen  
 ; APPLICANT: Xu, Yongyao  
 ; APPLICANT: Hoersch, Sebastian  
 ; APPLICANT: Monahan, John

```

? APPLICANT: Meyers, Rachel E.
? APPLICANT: Bast Jr., Robert C.
? APPLICANT: Hortobagyi, Gabriel N.
? APPLICANT: Pusztai, Laos
? APPLICANT: Meris, Funda
? APPLICANT: Sahin, Aysegul
? APPLICANT: Mills, Gordon B.
? TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
? TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
? FILE REFERENCE: MRI-018
? CURRENT APPLICATION NUMBER: US/10/177,293
? CURRENT FILING DATE: 2002-06-21
? PRIOR APPLICATION NUMBER: US 60/299,887
? PRIOR FILING DATE: 2001-06-21
? PRIOR APPLICATION NUMBER: US 60/331,572
? PRIOR FILING DATE: 2001-06-27
? PRIOR APPLICATION NUMBER: US 60/306,501
? PRIOR FILING DATE: 2001-07-18
? PRIOR APPLICATION NUMBER: US 60/325,002
? PRIOR FILING DATE: 2001-09-25
? PRIOR APPLICATION NUMBER: US 60/362,595
? PRIOR FILING DATE: 2002-03-05
? PRIOR APPLICATION NUMBER: US 60/xxx,xxx
? PRIOR FILING DATE: 2002-05-14
? NUMBER OF SEQ ID NOS: 506
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 45
? LENGTH: 999
? TYPE: PRT
? ORGANISM: Homo sapiens
? US/10/177,293-45

```

```

Query Match      100.0%; Score 5605; DB 15; Length 999;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 999; Conservative 3; Mismatches 0; Gaps 3

QY      1  MGVGRNRPGAAMAVLLLLLLLPLLJJAGAVPPGRGAAGPQEDVDECAQGLDCHXADA 60
DB      1  MGVGRNRPGAAMAVLLLLLLLPLLJJAGAVPPGRGAAGPQEDVDECAQGLDCHXADA 60

QY      61  LCQNTPTSYKSCRPYQGESRQCEDIDECNGLNGCVHDCNLINIGNVRCCTCFQGFMLA 120
DB      61  LCQNTPTSYKSCRPYQGESRQCEDIDECNGLNGCVHDCNLINIGNVRCCTCFQGFVLA 120

```

54: POAQRDSTPKEMFITVEFELETNQKENTASCD:SCIVKRTFKRLRKAIRTRKAVHREQ 600

```

Db      541  PCAGRPSTPKEMFIVPELETNQEVNTASCDLSIVVXRTKRLKKAIRTLKKAHREQ 600
Qy      601  FHLQLSGMNLDVAKKPPRTSERAQBSGCGQGAHAENOCVSCRAQTYDGAERERCILCPNG 660
Db      601  FHLQLSGMNLDVAKKPPRTSERAQBSGCGQGAHAENOCVSCRAQTYDGAERERCILCPNG 660
Qy      661  TFQNEEQMTCEPCPRPGNSGALKTPPEANMSECGGLCPQGEYSADGFAPCOLCALGTFQ 720
Db      661  TFQNEEQMTCEPCPRPGNSGALKTPPEANMSECGGLCPQGEYSADGFAPCOLCALGTFQ 720
Qy      721  PEAGRTSCPPCGGLATKQGATSFQDCETRVQCSGPHYNTTHRCIRCPVGTQPEFG 780
Db      721  PEAGRTSCPPCGGLATKQGATSFQDCETRVQCSGPHYNTTHRCIRCPVGTQPEFG 780
Qy      781  KNCVSCPGNTTDFDGSNTITCKNRRCGGELGDFGTGIESPNVPGNYPANTECTWTIN 840
Db      781  KNCVSCPGNTTDFDGSNTITCKNRRCGGELGDFGTGIESPNVPGNYPANTECTWTIN 840
Qy      841  PPKRRILIVVPEIFLPIEDDCGGLVLMKRTSSNSVTTTETCQYERPIAFTSRSKLW 900
Db      841  PPKRRILIVVPEIFLPIEDDCGGLVLMKRTSSNSVTTTETCQYERPIAFTSRSKLW 900
Qy      901  IQPKSNEGNSARGFQVPYVYDQYELIEDIVRGRLYASENHQELIKKXLIKALFDV 960
Db      901  IQPKSNEGNSARGFQVPYVYDQYELIEDIVRGRLYASENHQELIKKXLIKALFDV 960
Qy      961  LAHPQNYFKYTAQESREMPFRPSFIRLLRSKVSRLRPYK 999
Db      961  LAHPQNYFKYTAQESREMPFRPSFIRLLRSKVSRLRPYK 999

```

## RESULT 3

```

US-10-176-847-50
; Sequence 50, Application US/10176847
; Publication No. US20030068636A1
; GENERAL INFORMATION:
; APPLICANT: Veiby, Pelter Ole
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF BREAST
; TUMORS AND OVARIAN CANCER
; FILE REFERENCE: MRI-039
; CURRENT APPLICATION NUMBER: US/10/176,847
; CURRENT FILING DATE: 2002-06-21
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50
; LENGTH: 997
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-176-847-50

```

```

Query Match      99.6%; Score 5581; DB 15; Length 997;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 996; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
Qy      1  XGVAGNRPPGAWAV--LLLLLPPLLLLAGVPPGRGAAGPOEDVDECAQGLDCHADA 60
Db      1  XGVAGNRPPGAWAV--LLLLLPPLLLLAGVPPGRGAAGPOEDVDECAQGLDCHADA 58
Qy      61  LCQNTPTSYKSCCKPGYGGGROCEDIDCEGNELGVCVHDCLINIPGNVRCCTCFDGMILA 120
Db      59  LCQNTPTSYKSCCKPGYGGGROCEDIDCEGNELGVCVHDCLINIPGNVRCCTCFDGMILA 118
Qy      121  HDGHNCLDVECLENNGGCQHTCNVWMSYECCCKEGFFLSDNQHTCIHRSBEGLSCKMK 180
Db      119  HDGHNCLDVECLENNGGCQHTCNVWMSYECCCKEGFFLSDNQHTCIHRSBEGLSCKMK 178
Qy      181  DHGCSHICKAPRGSVACRCRPFELAKNQORDCILTNHNGGCGQSCDDTDAGPECSCH 240
Db      179  DHGCSHICKAPRGSVACRCRPFELAKNQORDCILTNHNGGCGQSCDDTDAGPECSCH 238
Qy      241  PQYNMHTDGRSCLEREDTVLEVTESNTTSVVDGDKVKRRLMETCAVNGGCDRTCKDT 300

```

```

Db      239  FQYKMTDGRSCLEREDTVLEVTESNTTSVVDGDKVKRRLMETCAVNGGCDRTCKDT 298
Qy      301  STGVHSCVPVGTFLQDLGKCTCKDIDECQTRNGGCDHFCKNTVGSFGCKGCKFKLLTDEK 360
Db      299  STGVHSCVPVGTFLQDLGKCTCKDIDECQTRNGGCDHFCKNTVGSFGCKGCKFKLLTDEK 358
Qy      361  SCQDVDECSLDRTCDHSCINHPGTAFACNCRVGLYGFTHCGDTNECSINNGGCGQCVN 420
Db      359  SCQDVDECSLDRTCDHSCINHPGTAFACNCRVGLYGFTHCGDTNECSINNGGCGQCVN 418
Qy      421  TVGSYECQCHPGKLNKXKDCVEVYKGLLPTSVPRVSLHCKGSGGDCGFLCHSGHIL 480
Db      419  TVGSYECQCHPGKLNKXKDCVEVYKGLLPTSVPRVSLHCKGSGGDCGFLCHSGHIL 478
Qy      481  SSDVTTTTSVTFKLNEGKCSLKNABLFPGLRPPALPEKHSSVKESFRVYVNLTCSSGKV 540
Db      479  SSDVTTTTSVTFKLNEGKCSLKNABLFPGLRPPALPEKHSSVKESFRVYVNLTCSSGKV 538
Qy      541  PGAPGRSTPKEMFIVPELETNQEVNTASCDLSIVVXRTKRLKKAIRTLKKAHREQ 600
Db      539  PGAPGRSTPKEMFIVPELETNQEVNTASCDLSIVVXRTKRLKKAIRTLKKAHREQ 598
Qy      601  FHLQLSGMNLDVAKKPPRTSERAQBSGCGQGAHAENOCVSCRAQTYDGAERERCILCPNG 660
Db      599  FHLQLSGMNLDVAKKPPRTSERAQBSGCGQGAHAENOCVSCRAQTYDGAERERCILCPNG 658
Qy      661  TFQNEEQMTCEPCPRPGNSGALKTPPEANMSECGGLCPQGEYSADGFAPCOLCALGTFQ 720
Db      659  TFQNEEQMTCEPCPRPGNSGALKTPPEANMSECGGLCPQGEYSADGFAPCOLCALGTFQ 718
Qy      721  PEAGRTSCPPCGGLATKQGATSFQDCETRVQCSGPHYNTTHRCIRCPVGTQPEFG 780
Db      719  PEAGRTSCPPCGGLATKQGATSFQDCETRVQCSGPHYNTTHRCIRCPVGTQPEFG 778
Qy      781  KNCVSCPGNTTDFDGSNTITCKNRRCGGELGDFGTGIESPNVPGNYPANTECTWTIN 840
Db      779  KNCVSCPGNTTDFDGSNTITCKNRRCGGELGDFGTGIESPNVPGNYPANTECTWTIN 838
Qy      841  PPKRRILIVVPEIFLPIEDDCGGLVLMKRTSSNSVTTTETCQYERPIAFTSRSKLW 900
Db      839  PPKRRILIVVPEIFLPIEDDCGGLVLMKRTSSNSVTTTETCQYERPIAFTSRSKLW 898
Qy      901  IQPKSNEGNSARGFQVPYVYDQYELIEDIVRGRLYASENHQELIKKXLIKALFDV 960
Db      899  IQPKSNEGNSARGFQVPYVYDQYELIEDIVRGRLYASENHQELIKKXLIKALFDV 958
Qy      961  LAHPQNYFKYTAQESREMPFRPSFIRLLRSKVSRLRPYK 999
Db      959  LAHPQNYFKYTAQESREMPFRPSFIRLLRSKVSRLRPYK 997

```

## RESULT 4

```

US-09-855-824-5
; Sequence 5, Application US/09655824
; Publication No. US20030166048A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; METHODS OF IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; BREAST TUMORS AND OVARIAN CANCER
; FILE REFERENCE: US/09/855,824
; CURRENT APPLICATION NUMBER: US/09/855,824
; CURRENT FILING DATE: 2001-05-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 964
; TYPE: PRT
; ORGANISM: Human
US-09-855-824-5

```

Query Match: 96.8%; Score 5428; DB 12; Length 964;



```

QY 721 PEAGRTSCPCGGGLATKHGATSFQDCETRVQCSQPGHFYNTTTHRCIRCPVGTQYQEPFG 780
Db 729 PDVGRITSCLSCGGGLPTKHGATSFQDCETRVQCSQPGHFYNTTTHRCIRCPVGTQYQEPFG 776
QY 781 KNNVSCPGNTTDFDGSNTITCKNRRCGGELGDFGTGIESPNYPGNYPANTECTWTIN 840
Db 779 KNNVSCPGNTTDFDGSNTITCKNRKCGELGDFGTGIESPNYPGNYPANTECTWTIN 838
QY 841 PPKRRILLIIVPEIFLPIEDDCGYLVMRKTSSNSVTTYETCOTYERPIAFTSSKKLM 900
Db 839 PPKRRILLIIVPEIFLPIEDDCGYLVMRKTSSNSVTTYETCOTYERPIAFTSSKKLM 898
QY 901 IQFKSNEGNSARGFQVPYVYTDYDYOELIEDIVRDCRLYASENHQELIKKKLILKALFDV 960
Db 899 IQFKSNEGNSARGFQVPYVYTDYDYOELIEDIVRDCRLYASENHQELIKKKLILKALFDV 958
QY 961 LAHPQNYFKYTAQESREMFPRPSIRLLRSKVSFLRPYK 999
Db 959 LAHPQNYFKYTAQESREMFPRPSIRLLRSKVSFLRPYK 997

RESULT 6
US-09-855-824-6
; Sequence 6, Application: US/09855824
; Publication No. US20030166048A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001233
; CURRENT APPLICATION NUMBER: US/09/855,824
; CURRENT FILING DATE: 2001-05-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 957
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-855-824-6

Query Match 88.08; Score 4935; DB 12; Length 957;
Best Local Similarity 90.21; Pred. No. 0;
Matches 862; Conservative 46; Mismatches 48; Indels 0; Gaps 0;

QY 44 EDVDECAQGLDDCHADALCQNTPTSYKCSCKPGYQGEGRQCEDIDBCGNELNGGVHDCI 103
Db 2 EDVDECAQGLDDCHADALCQNTPTSYKCSCKPGYQGEGRQCEDIDBCGNELNGGVHDCI 61
QY 104 NIPGNVRCCTCFDGFMLADHGNCLDVBDELENNGGQHTCVNVMGSYECCCKEGFPLSDN 163
Db 62 NIPGNVRCCTCFDGFMLADHGNCLDVBDELENNGGQHTCVNVMGSYECCCKEGFPLSDN 121
QY 164 QHTCIHRSSEGLSCMNKDHGCHICKAPRGSVACRCRPGFELAKNQRDCILTCNHGNG 223
Db 122 QHTCIHRSSEGLSCMNKDHGCHICKAPRGSVACRCRPGFELAKNQRDCILTCNHGNG 181
QY 224 QHSCDDTADGPECSCHPQYKMTDGRSCLEREDTVLEVTESNTTSVWDGDKRVKRRLLM 283
Db 182 QHSCDDTADGPECSCHPQYKMTDGRSCLEREDTVLEVTESNTTSVWDGDKRVKRRLLM 241
QY 284 ETCVAVNGGCDRTCKDTSTGVHSCVPGFTLQDGGTKCIDBCQTRNGGCDHFCKNIYV 343
Db 242 ETCVAVNGGCDRTCKDTSTGVHSCVPGFTLQDGGTKCIDBCQTRNGGCDHFCKNIYV 302
QY 344 SFDGCKKGFLLTDEKSCQDVDECSLDRTCDHSCINHPHTFACACNRYTYLYGTGCD 403
Db 302 SFDGCKKGFLLTDEKSCQDVDECSLDRTCDHSCINHPHTFACACNRYTYLYGTGCD 362
QY 404 TNECSINNGGQQCVNTVSGYCCQHPGYKHLHNNKQCVKGLLPTSVSRVSLRQK 463
Db 362 TNECSINNGGQQCVNTVSGYCCQHPGYKHLHNNKQCVKGLLPTSVSRVSLRQK 421

```

```

QY 464 SGGDGCPLRCSSGCHLSSDVTTIRTSVTFXNKGKSLKNSLEPECSRPALPEKHSSV 523
Db 422 SGGDGCPLRCSSGCHLSSDVTTIRTSVTFXNKGKSLKNSLEPECSRPALPEKHSSV 481
QY 524 KESFRVNLTCSSGKQVPGAPRPTPKEMFTTVFPELETNOKEVTASCDLSCYKRTK 583
Db 482 KESFQVNLTCSSGKQVPGALGRJNAPKEMFTTVFERETYEKVTASCNLSCVYKRTK 541
QY 584 RLKAIPTLKAHREDFHLOLSGNLIDVAKPPPTSERQAESCGVGHAEHNCQVSGRA 643
Db 542 RLKAIPTLKAHREDFHLOLSGNLIDVAKPPPTSERQAESCGVGHAEHNCQVSGRA 601
QY 644 GTTYDGARERCILCPNGTTFQNEEGQMTCEPPRPNNGSALKTPEAMNMSCEGGLCQPEY 703
Db 602 GTTYDGARERCILCPNGTTFQNEEGQMTCEPPRPNNGSALKTPEAMNMSCEGGLCQPEY 661
QY 704 SAGFAPCQCALGTPOPEAGRTSCPPCGGGLATKHGATSFQDCETRVQCSQPGHFYNT 763
Db 662 SAGFAPCQCALGTPOPEAGRTSCPPCGGGLATKHGATSFQDCETRVQCSQPGHFYNT 721
QY 764 THRCIRCPVGTQYQEPFGKNNVSCPGNTTDFDGSNTITCKNRRCGGELGDFGTGIESP 823
Db 722 THRCIRCPVGTQYQEPFGKNNVSCPGNTTDFDGSNTITCKNRRCGGELGDFGTGIESP 781
QY 824 NYPGNYPANTECTWTINPPPKRRILLIIVPEIFLPIEDDCGYLVMRKTSSNSVTTYETC 883
Db 782 NYPGNYPANTECTWTINPPPKRRILLIIVPEIFLPIEDDCGYLVMRKTSSNSVTTYETC 841
QY 884 QTYERPIAFTSSKKLMQFKSNEGNSARGFQVPYVYTDYDYOELIEDIVRDCRLYASEN 943
Db 842 QTYERPIAFTSSKKLMQFKSNEGNSARGFQVPYVYTDYDYOELIEDIVRDCRLYASEN 901
QY 944 HQEILKDKKLILKALFDVLAHPQNYFKYTAQESREMFPRPSIRLLRSKVSFLRPYK 999
Db 902 HQEILKDKKLILKALFDVLAHPQNYFKYTAQESREMFPRPSIRLLRSKVSFLRPYK 957

RESULT 7
US-09-855-824-2
; Sequence 2, Application: US/09855824
; Publication No. US20030166048A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL001233
; CURRENT APPLICATION NUMBER: US/09/855,824
; CURRENT FILING DATE: 2001-05-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 988
; TYPE: PRT
; ORGANISM: Human
US-09-855-824-2

Query Match 65.6%; Score 3675.5; DB 12; Length 988;
Best Local Similarity 62.4%; Pred. No. 14e-270;
Matches 654; Conservative 113; Mismatches 162; Indels 119; Gaps 14;

QY 10 GAA---WAVLLIILPPLLLLAGAVPPGGRGAAG-----POEDVDSCAQLDQCHADAL 61
Db 2 GAAAVWHLICVLLAL-----GTRGLAGSGGLPGSDVDYDECSGSDCHIDAL 49
QY 62 QNTPTSYKCSCKPGYQGEGRQCEDIDECGNE-LKSGCVHDCIN:PGNVRTCTCFGFMCLA 120
Db 50 QNTPTSYKCSCKPGYKGEKQCEIDEENDYNGGCVHCEINIPGNVYRCTCFGFMCLA 109
QY 121 HDGHNCLDVBDELENNGGQHTCVNVMGSYECCCKEGFPLSDNHTC:HRSEEGJSCMNK 180
Db 110 HDGHNCLDVBDECDNNGGQHTCVNVMGSYECCCKEGFPLSDNHTC:HRSEEGJSCMNK 169

```

181 DHGCSHICKEAPRGVACBCEPFLAKNORCILTCHNGGCGHSCDDTADGPECSCH 240  
 170 DHGCAHICRETAKGVACDCRFGFLAQKQKCTLTCHNGGCGHSCDDTADGPECSCH 229  
 241 POYKXHTDGRSLEREDTVLEVTESNTTSDVDDGKVRKRLMETCAVNNGGCORTCKDT 300  
 230 QKVALHSDGRTCI-----ETCAVNNGGCORTCKDT 259  
 301 STGVHSCVPVGTQLDQKTKDIDECQTRNGGCHFCNIVGSDCCCKKQKFLITDEK 360  
 260 ATGVHSCVPVGTQLDQKTKDIDECQTRNGGCHFCNIVGSDCCCKKQKFLITDEK 319  
 361 SCQVDECSLDRTCDHSCNHPOTFACACNRGTYLGFTHCGGTNECSINNGGCGQCVN 420  
 320 TCQDIDECFERTCDHICNSPGSFQCLCHRGYLYGTHCGDVDECSNNGSCDQCGVN 379  
 421 TVGSYECQCHPGYKHLHNNKCCVEV-KGLLPTSVSPRVSLHCGKGGGGGFLRC----- 474  
 380 TKGSYECVCPGPRRLHNNKDCVETGCKLSRAKTSAPRAQLSCSKAGGVESCF-SCPAHTL 439  
 475 -----HSGIHL-----SDVTITRTSVTFKLNKCKSL 502  
 440 FVPDSSENSVYVLCGVPPGQKALQKNGTSSGLGPGSCSDAPTPIKOKARFKIRDAKCHL 499  
 503 KNAELFPEGLRPALEKHSVSKESPR-----YVNLTCSSGKQVPGAPGRSTPKE 552  
 500 R-----PHSQARAKETARQPLLDHCHVTFTLKCDSKK--RRGRKSPSKE 544  
 553 M-FITVEFELETNQKEVTASCDLSICVIRKTEKRLKAKAIRTLRKAVHREOFHQLSGXND 611  
 545 VSHITAEFELETKYBEASDTCEADCLRKAEQSLQAIAKTLKRSIGRQFYVQSGTEYE 604  
 612 VAKKPRTSERQAESCGVQGHAEQVSCRACTYDGAERECILCPNGTQNEGGQTC 671  
 605 VAQRPAKALEGGG-ACGAGQVLDQSKVACGPGTHFGGELGQCVSCMPGTQDMEGQJSC 663  
 672 EPCPRGNSGALKTPKAMNMSCGGLCPQGEYSADGAPCALCGLTQPAAGRTSCFPC 731  
 664 TPCP---SSDGLGLPGARNVSECGGCGSPGFSADGFKPCQACPVGTQPEBRTGCEPC 720  
 732 GGLATKHQCATSFQDCETRVQCSGPHFYNTTHRCIRCPVGTQPEFGKNCVSCPGNT 791  
 721 GGLLTKEGTTSFQDCEAKVCSGPHYNTTHRCIRCPVGTQPEFGQNHCTCPGNT 780  
 792 TTDFDGSNTITCKNRRCGSELGDFGTGYIESPNYPGNYPANTECTWTINPPKRIILVW 851  
 781 STDFDGSNTVTHCKNCHGCGELGDFGTGYIESPNYPGDPYANAEVWHIAPPKRIILVW 840  
 852 PEIFLPIDDCGDLVLRKTSNSNVTYETCTYERPIAFTSRSKK-LWOFKSNEGNSA 911  
 841 PEIFLPIDECGDLVLRKTSASPTSIITYETCTYERPIAFTSRSKKLWOFKSNEGNSG 900  
 912 RGFQVYVYDYEDYQELIEDIVDRGLYASENHQILKDKK-LKALFDVLAPHPQNYFYKT 971  
 901 RGFQVYVYDYEDYQELIEDIVDRGLYASENHQILKDKK-LKALFDVLAPHPQNYFYKT 960  
 972 AQESREMPFRSIRLRSKVSRLRPYK 999  
 961 AQESREMPFRSIRLRSKVSRLRPYK 988

RESULT 8

US-10-195-142-2

Sequence 2. Application US/10195:142

Publication No. US20030036163A1

GENERAL INFORMATION:

APPLICANT: Myriad Genetics, Inc.

APPLICANT: Wettstein, Daniel Albert

APPLICANT: Mauck, Kimberly A.

TITLE OF INVENTION: NOVEL PN9826 NUCLEIC ACIDS AND USE THEREOF

FILE REFERENCE: 1051.02

CURRENT APPLICATION NUMBER: US/10-195,142

CURRENT FILING DATE: 2002-07-10

PRIOR APPLICATION NUMBER: US 60/304,323  
 PRIOR FILING DATE: 2001-07-10  
 NUMBER OF SEQ ID NOS: 39  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 2  
 LENGTH: 988  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-195-142-2

Query Match 65.6%; Score 3675.5; DB 15; Length 988;  
 Best Local Similarity 62.4%; Pred. No. 14e-270;  
 Matches 654; Conservative 113; Mismatches 162; Indels 115; Gaps 14;

QY 10 GAA---WAVJLLLLLPULLLAGAVPPGRGGAAG-----PQEDVDEACAGLDDCHADAL 61  
 DB 2 GAAARWHLVCLLAL-----GTRGLAGGSLPGSVVDVDECEGTDDCHIDAI 49  
 QY 62 CONTPTSYKSCSKPGYGGEGROCEDIDECGNE-LJGGCVHDCNLINPGNYRCTCFDGFMLA 120  
 DB 50 CONTPKSYKCLCKPGYKGEKQCEDIDECENDYNGGCVHECINPGNTRCTCFDGFMLA 109  
 QY 121 HDGHNCLVDDECLENNCGQHTCVWVMGYVRCCKEGFFLSPNQHCTCHRSBEGSLCKMK 180  
 DB 110 HDGHNCLVDDECLENNCGQHTCVWVMGYVRCCKEGFFLSPNQHCTCHRSBEGSLCKMK 169  
 QY 181 DHGCSHICKEAPRGVACBCEPFLAKNORCILTCHNGGCGHSCDDTADGPECSCH 240  
 DB 170 DHGCAHICRETAKGVACDCRFGFLAQKQKCTLTCHNGGCGHSCDDTADGPECSCH 229  
 QY 241 POYKXHTDGRSLEREDTVLEVTESNTTSDVDDGKVRKRLMETCAVNNGGCORTCKDT 300  
 DB 230 QKVALHSDGRTCI-----ETCAVNNGGCORTCKDT 259  
 QY 301 STGVHSCVPVGTQLDQKTKDIDECQTRNGGCHFCNIVGSDCCCKKQKFLITDEK 360  
 DB 260 ATGVHSCVPVGTQLDQKTKDIDECQTRNGGCHFCNIVGSDCCCKKQKFLITDEK 319  
 QY 361 SCQVDECSLDRTCDHSCNHPOTFACACNRGTYLGFTHCGGTNECSINNGGCGQCVN 420  
 DB 320 TCQDIDECFERTCDHICNSPGSFQCLCHRGYLYGTHCGDVDECSNNGSCDQCGVN 379  
 QY 421 TVGSYECQCHPGYKHLHNNKCCVEV-KGLLPTSVSPRVSLHCGKGGGGGFLRC----- 474  
 DB 380 TKGSYECVCPGPRRLHNNKDCVETGCKLSRAKTSAPRAQLSCSKAGGVESCF-SCPAHTL 439  
 QY 475 -----HSGIHL-----SDVTITRTSVTFKLNKCKSL 502  
 DB 440 FVPDSSENSVYVLCGVPPGQKALQKNGTSSGLGPGSCSDAPTPIKOKARFKIRDAKCHL 499  
 QY 503 KNAELFPEGLRPALEKHSVSKESPR-----YVNLTCSSGKQVPGAPGRSTPKE 552  
 DB 500 R-----PHSQARAKETARQPLLDHCHVTFTLKCDSKK--RRGRKSPSKE 544  
 QY 553 M-FITVEFELETNQKEVTASCDLSICVIRKTEKRLKAKAIRTLRKAVHREOFHQLSGXND 611  
 DB 545 VSHITAEFELETKYBEASDTCEADCLRKAEQSLQAIAKTLKRSIGRQFYVQSGTEYE 604  
 QY 612 VAKKPRTSERQAESCGVQGHAEQVSCRACTYDGAERECILCPNGTQNEGGQTC 671  
 DB 605 VAQRPAKALEGGG-ACGAGQVLDQSKVACGPGTHFGGELGQCVSCMPGTQDMEGQJSC 663  
 QY 672 EPCPRGNSGALKTPKAMNMSCGGLCPQGEYSADGAPCALCGLTQPAAGRTSCFPC 731  
 DB 664 TPCP---SSDGLGLPGARNVSECGGCGSPGFSADGFKPCQACPVGTQPEBRTGCEPC 720  
 QY 732 GGLATKHQCATSFQDCETRVQCSGPHFYNTTHRCIRCPVGTQPEFGKNCVSCPGNT 791  
 DB 721 GGLLTKEGTTSFQDCEAKVCSGPHYNTTHRCIRCPVGTQPEFGQNHCTCPGNT 780  
 QY 792 TTDFDGSNTITCKNRRCGSELGDFGTGYIESPNYPGNYPANTECTWTINPPKRIILVW 851  
 DB 781 STDFDGSNTVTHCKNCHGCGELGDFGTGYIESPNYPGDPYANAEVWHIAPPKRIILVW 840  
 QY 852 PEIFLPIDDCGDLVLRKTSNSNVTYETCTYERPIAFTSRSKK-LWOFKSNEGNSA 911  
 DB 841 PEIFLPIDECGDLVLRKTSASPTSIITYETCTYERPIAFTSRSKKLWOFKSNEGNSG 900  
 QY 912 RGFQVYVYDYEDYQELIEDIVDRGLYASENHQILKDKK-LKALFDVLAPHPQNYFYKT 971  
 DB 901 RGFQVYVYDYEDYQELIEDIVDRGLYASENHQILKDKK-LKALFDVLAPHPQNYFYKT 960  
 QY 972 AQESREMPFRSIRLRSKVSRLRPYK 999  
 DB 961 AQESREMPFRSIRLRSKVSRLRPYK 988



```
QY 852 PEIPLIEDDCGVLMRKTSSTSSNVTTTCTQYERPIAFTSRKLLWQKSNRNSA 912
DB 841 PEIPLIEDDCGVLMRKTSSTSSNVTTTCTQYERPIAFTSRKLLWQKSNRNSG 900
QY 912 RGFQVPYVYDDEYQELIEDVDRGRLYASENHQELKDKKLIKALFDVLAHFQNYFKYT 971
DB 901 KGFQVPYVYDDEYQELIEDVDRGRLYASENHQELKDKKLIKALFDVLAHFQNYFKYT 960
QY 972 AQESREMPFRSFIKLLRSKVSRLRPYK 999
DB 961 AQESREMPFRSFIKLLRSKVSRLRPYK 988

RESULT 9
US-09-898-570-14
; Sequence 14, Application US/09898570
; Patent No. US20020123612A1
; GENERAL INFORMATION:
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND
; TITLE OF INVENTION: METHODS OF USING THE SAME
; FILE REFERENCE: 15966-776CIP
; CURRENT APPLICATION NUMBER: US/09/898,570
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/198,293
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/198,645
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/210,809
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/199,476
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,025
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/224,610
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/200,024
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/199,880
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/218,591
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/271,814
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 60/215,855
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 09/839,446
; PRIOR FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 14
; LENGTH: 974
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: POLYX
; OTHER INFORMATION: 297832_B_1
US-09-898-570-14

Query Match 61.0%; Score 3417.5; DB 10; Length 974;
Best Local Similarity 60.1%; Pred. No. 5.7e-251;
Matches 596; Conservative 140; Mismatches 219; Indels 37; Gaps 9;

QY 22 LPPLLLAGVPPGGRAGQEDVDECAQGLDDCHADALCONTPTSYKSCXKPGYQEG 81
DB 6 VPEGLLVLLVHARAQYKAAQDVDECVETDNCNCHDAICONTPSYKICKSGYTG 65
QY 92 RCEIDECGELNGCGVHDCINIPGNRYCTCFDGFMLAHGHNCLDVBCLNNGSCQH 141
```

```
DE 66 KHKCQVDECEDEEDNAGCVHDCVNI:PGNYRCTCYDGFHLADGHNCNCLDVECAEGNGGCOQ 125
QY 142 TCNVNMGSYCCCKBESFFILSDNCHTCHIRSEBGLSCMNKDHGCSHICKAPRGSVACER 201
DB 126 SCNVNMGSYCHREGFFILSDNCHTCHIRSEBGLSCMNKDHGCSHICKAPRGSVACER 185
QY 202 PGFELAKNGSCCLTGNHNGGCGHSCDDTADQPECSCHPOYKVKHTDGRSCILERSDVTLE 261
DB 186 PGFELAKNGSCCLTGNHNGGCGHSCDDTADQPECSCHPOYKVKHTDGRSCILERSDVTLE 237
QY 262 VTESNTTSVVDGSKVRKRLJ-----LVETCAVNNGGCCTCTKOTSFGVHCSCPVFTLOJ 316
DB 238 -----GERRLECHPTCAVSNETCAVNNGGCDSKCHDAATGVHCTCPVGFMLQP 286
QY 317 DGTCKDIDEQOTRNGGCHFCNIVGSEFCCKKFKLLTDEKSCODVDESLORTCOH 376
DB 287 DRKCKDIDEQOTRNGGCHFCNIVGSEFCCKKFKLLTDEKSCODVDESLORTCOH 346
QY 377 SCINHPGTFAACNRYGTYLGFTHCGDTNECSINNGGCGQCVNTVGSYECOCCHPOY-KJ 435
DB 347 ICYNTPGSFQCLCHRGYLLVGI:THCGDVDECSINRGCGRFGCINTPGSYOCTCPAQOGR 406
QY 436 HWKKKDCVEKGLLPTSVSPRSLHCKSGGGGCGFLRCHSGIHL--SSDVTTIRISVTF 493
DB 407 HWNGKDCTEPLKCGSPGASKAMLSNRSCKKDTCAITCPSRARFLPEAAVLISKORASF 466
QY 494 KLNKGC--SLKNAELFPGLRLPALPEKHSSVKESPRYVNLTCSSGKQVPGAPRSTPK 551
DB 467 KIKDAKRLH:RNKGTTEAGRTTGGGAPCSCQVTFIHLKDCDSRKSKGRARTPPCK 526
QY 552 EMP+TVFEFELEINQKEVTASCDLSIVKRTKRLKARTL:RKAVHRSQFHLQSLGML 610
DB 527 EVTRLELEAEVPAEETTASQGLPCLRQRMERRLKGSLKLRKSLNQRELLRLAGLDY 586
QY 611 DVAKKPRPTSERQA---ESCGVGQGHAEHQVSCRAGTYVDGARERCILCPNGTCQNEEG 667
DB 587 EAHKPGLVAGERAEPNVEPCQHRAGTKVSCPPGTYHGO:BOCVPCFAGTFOEREG 646
QY 668 QMTCEPCPRPNSGALKTPKAMNSECGLCQPGEYSADGFAPCQLCALGTTPQEAGRTS 727
DB 647 QLSCLCPGSDAHGFL--GATNVTTICAGCPCPGQSVDFKPCQCPRGTYOPEAGRTL 703
QY 728 CFPGGGLATKHQNTSFQDCETRVQSGPHFNTTHCIRCPVGTYOPEGKNVCSC 787
DB 704 CFPGGGLATKHQNTSFQDCETRVQSGPHFNTTHCIRCPVGTYOPEGKNVCSC 763
QY 788 PGNITTFDGSNITCKNRRCGGELGDFTYIESPNYPONYPANTECTWT:NPPEKRI 847
DB 764 PGNITTFDGSNITCKNRRCGGELGDFTYIESPNYPONYPANTECTWT:NPPEKRI 823
QY 848 LIVVPEIFLPIEDDCGVLMRKTSSTSSNVTTTCTQYERPIAFTSRKLLWQKSNRNS 907
DB 824 LIVVPEIFLPIEDDCGVLMRKTSSTSSNVTTTCTQYERPIAFTSRKLLWQKSNRNS 883
QY 908 GNSARGQVYVYDDEYQELIEDVDRGRLYASENHQELKDKKLIKALFDVLAHFQNY 967
DB 884 ANSARGQVYVYDDEYQELIEDVDRGRLYASENHQELKDKKLIKALFDVLAHFQNY 943
QY 968 FKYTAQSRMFPFRSFIKLLRSKVSRLRPYK 999
DB 944 FKYT-EKHKEMLPKSPFKLLRSKVSRLRPYK 974

RESULT 10
US-09-899-446-14
; Sequence 14, Application US/09839446
; Publication No. US20030050232A1
; GENERAL INFORMATION:
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND
```

## TITLE OF INVENTION: METHODS OF USING THE SAME

FILE REFERENCE: 15966-776

CURRENT APPLICATION NUMBER: US/09/839,446

CURRENT FILING DATE: 2001-04-19

PRIOR APPLICATION NUMBER: 60/198,293

PRIOR FILING DATE: 2000-04-19

PRIOR APPLICATION NUMBER: 60/198,645

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: 60/210,909

PRIOR FILING DATE: 2000-06-09

PRIOR APPLICATION NUMBER: 60/199,476

PRIOR FILING DATE: 2000-04-26

PRIOR APPLICATION NUMBER: 60/200,025

PRIOR FILING DATE: 2000-04-26

PRIOR APPLICATION NUMBER: 60/224,610

PRIOR FILING DATE: 2000-08-11

PRIOR APPLICATION NUMBER: 60/200,024

PRIOR FILING DATE: 2000-04-26

PRIOR APPLICATION NUMBER: 60/199,880

PRIOR FILING DATE: 2000-04-26

PRIOR APPLICATION NUMBER: 60/218,591

PRIOR FILING DATE: 2000-07-17

PRIOR APPLICATION NUMBER: 60/271,814

PRIOR FILING DATE: 2001-02-27

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 14

LENGTH: 974

TYPE: PRT

ORGANISM: Unknown Organism

FEATURE:

OTHER INFORMATION: Description of Unknown Organism: POLYX

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

US-09-839-446-14

## Query Match

Best Local Similarity: 60.4%; Score 3477.5; DB 11; Length 974;

Matches 596; Conservative 140; Mismatches 219; Indels 37; Gaps 9;

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

## Query Match

Best Local Similarity: 56.3%; Pred. No. 2; Length 993;

Matches 556; Conservative 136; Mismatches 212; Indels 78; Gaps 9;

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36

US-10-239-663-36



```

QY 377 SCINHPGTACACNSGYTLGFTHCGDNEGSINNGCCQYCVANTVSGYECQCHPGY-KL 435
DB 347 ICVNTPGSPQCLCHRGYLLYSITTHCGDVECSINNGGRCFGCINTPGSYCCTCPAGQGR 406
QY 436 HMKKDCVEYKGLLPTSVSPRYSJHCGSGGGDCCFLAC----- 474
DB 407 HMKKDCVEYKGLLPTSVSPRYSJHCGSGGGDCCFLAC----- 474
QY 475 -----HSGIHLSSD-----VTTIRTSVTFKLNKGC--SLKNAEAFPEGJRFAL 516
DB 467 SPRAAPARAGHNGNSTNSHCHAEAVLSIKQASAKIKDANCRLLRNKKTTEAGRTG 526
QY 517 PEKSSSVKESFRYYNLTSSGKQVPGAPRSTPKEMF-ITVEFELETOKEVTAACDLS 575
DB 527 PGAPRCECCQVTFIHLKDCSSRKGRARATPPGKEVTRLTLEAEVRAEETTAACGLP 586
QY 576 CIVKREKRLKAIATLBRKAVHREDFHLQLSGMNLDVAKKPPRTSERQA---ESGCVGG 632
DB 587 CTRQMERRLKGSJMKLRKSIQCFFLLRLAGLDYEJAHKGLVAGERAEMECRPGCH 646
QY 633 HAENOCVCRAGTYVDGARERCILCPNGTFQNEBQMTCEPCPRPGNSGALKTPEAMNKS 692
DB 647 RAGTKCVSCPGCTTYHGGTEQCVCPCAGTFQREBQSLSDLCPSDAHGPL---GATVVT 703
QY 633 ECGGLCQPGEVSADGFAPRCULCALGTFQPEAGRTSCFPGGGLATKQGSATSPDDCTRV 752
DB 764 TCAGCCPPGQHSVDFKPCQCPGRTGQPEAGRTLCFPGGGGLTKHGAISFDDCDTKV 763
QY 753 QCSPGHFNITTHRCIRCPVGTGQPEFGKNQVSGRNTTDFDGSNTITQCKKRRGCGE 812
DB 764 QCSPGHFNITTHRCIRCPVGTGQPEFGKNQVSGRNTTDFDGSNTITQCKKRRGCGE 823
QY 813 LGDTGYIESPNYPNPANTECTWTINPPKRLILIVPELFLPIEDCGDYLMKXTS 872
DB 824 LGFTGYIESPNYPNPANTECTWTINPPKRLILIVPELFLPIEDCGDYLMKXTS 883
QY 873 SSNVTYETCTYETRPATFTSRKMLIQFKSNGSARGQVYVYDDEYQELIIDI 932
DB 884 SPSITTYETCTYETRPATFTSRKMLINFTTSEANSARGQVYVYDDEYQELIIDI 943
QY 933 VDDGLVASENHQELIKOKKIKALFDVLAPHQNYFKYTAQESREMPFRSFLRSKVS 992
DB 944 VDDGLVASENHQELIKOKKIKALFDVLAPHQNYFKYTAQESREMPFRSFLRSKVS 1002
QY 993 RELRPYK 999
DB 1003 RELRPYK 1009

```

```

PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: 60/224,610
PRIOR FILING DATE: 2000-08-11
PRIOR APPLICATION NUMBER: 60/200,024
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: 60/199,880
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: 60/218,591
PRIOR FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 60/271,814
NUMBER OF SEQ ID NOS: 49
SOFTWARE: Patent Ver. 2.1
SEQ ID NO: 16
LENGTH: 1009
TYPE: PRT
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: POLYX
OTHER INFORMATION: G355056-04
US-09-839-446-16

Query Match 60.6% Score 3394; DE 11; Length 1009;
Best Local Similarity 58.1%; Pred. No. 3 6e-249;
Matches 597; Conservative 142; Mismatches 216; Indels 72; Gaps 10;

QY 22 LPELLLJAGAVPFGRRAPAPCEDEVEDCAQGLDDCHADALCONTFSTYKSGCKFYQGS 81
DB 6 VPEGLCLVLLVHARAQYKAAQDVDECEVTGNCHIDAICQVTPRSYKICIKSGYTGDS 65
QY 82 ROCEJIDECNE-NGCCVHDCJNIPENYECTGDSKYLADHGNICDVDECCJENNNGGCH 141
DB 66 RKQKCVDECEBENACVADCVNIPENYCTCYDGFHLADHGNICDVDECCJENNNGGCH 125
QY 142 TCYNMMSYECCCKESGFPFLSDNHTTIFHSEESJSGMNDHGSNICKAPRGSVACER 201
DB 126 SCVNMMSYCHCRBGFLLSDNHTTIOREBEMNMNKHGCAHICRETPKGGJACECR 165
QY 202 PGFELAKXPRCITLTNHSNGSCSHSCDPTADBPESCHPQYKATDGSGLEREDTVJE 261
DB 186 PGFELTNQPRCKLTCTYNSGGCQHTCEDTEGPRGCHIKFVLTIDGKTCI----- 237
QY 262 VTESNTSVDDGXRYKRR-----LMEICAVNNGSCDPTCKDTSTGYVCSGCVGFTLO 316
DB 238 -----GRRLEQHTPTQAVSNBTQAVNNGSCDSCKCPAALGVHCTGCVGFTLO 286
QY 317 DGKTCXIDECQTRNGGCDHFCXKIVGSPDCCGCKGFKLLTDEKSCQDVDECSJDRTCDH 376
DB 287 DRKTCXIDECRANNGGCHICRNTVGSFECCKGYKLLINRNCQDIDEBRFDRTCDH 346
QY 377 SCINHPGTACACNSGYTLGFTHCGDNEGSINNGCCQYCVANTVSGYECQCHPGY-KL 435
DB 347 ICVNTPGSPQCLCHRGYLLYSITTHCGDVECSINNGGRCFGCINTPGSYCCTCPAGQGR 406
QY 436 HMKKDCVEYKGLLPTSVSPRYSJHCGSGGGDCCFLAC----- 474
DB 407 HMKKDCVEYKGLLPTSVSPRYSJHCGSGGGDCCFLAC----- 474
QY 475 -----HSGIHLSSD-----VTTIRTSVTFKLNKGC--SLKNAEAFPEGJRFAL 516
DB 467 SPRAAPARAGHNGNSTNSHCHAEAVLSIKQASAFIKQAKKSLHLRNKKTTEAGRTG 526
QY 517 PEKSSSVKESFRYYNLTSSGKQVPGAPRSTPKEMF-ITVEFELETOKEVTAACDLS 575
DB 527 PGAPRCECCQVTFIHLKDCSSRKGRARATPPGKEVTRLTLEAEVRAEETTAACGLP 586
QY 576 CIVKREKRLKAIATLBRKAVHREDFHLQLSGMNLDVAKKPPRTSERQA---ESGCVGG 632
DB 587 CTRQMERRLKGSJMKLRKSIQCFFLLRLAGLDYEJAHKGLVAGERAEMECRPGCH 646
QY 633 HAENOCVCRAGTYVDGARERCILCPNGTFQNEBQMTCEPCPRPGNSGALKTPEAMNKS 692
DB 647 RAGTKCVSCPGCTTYHGGTEQCVCPCAGTFQREBQSLSDLCPSDAHGPL---GATVVT 703

```

QY 693 ECGGLCCPGEVSADGAPACQCALGTFQPEAGRTSCPPCGGLATKHOGATSPQCESTRY 752  
 DB 704 TCAQCPCPGGHSVGVGFQCPQPRGTQPEAGRTLCPPCGGLTKKHEGA:SPQDCDTRV 753  
 QY 753 QCSGPHFYNTTTHRCIRCPVGTQPEFGKNNVCSPNNTTTFDGSNTITQCKNRPCCGE 812  
 DB 764 QCSGPHYNTTTHRCIRCMKMSYQPDPRQNCRCRPN:STDPFGSISVACNRCGGE 823  
 QY 813 LGDFGTYESPNYGNYPANTECTWTINPPKRRILIVPEILPIEDDCGCVLVMKTS 872  
 DB 824 LGSEFGYIESPNYGNYPAGVEICMINPPKRRILIVPEILF:PSSEDCGCVLVMKNS 883  
 QY 873 SSNSVTYETQTERPIAFTSRKSLMIOFKSNEGSARGFQVRYTYEDYQELLEDI 932  
 DB 884 SPSSITVETQTERPIAFTSRKSLMINKTSSEANSARGFQVRYTYEDYQELLEDI 943  
 QY 933 VRDRLYASENHQELIKDKLILK:FDVLAPQVRYTYAQESREMPFRAFIILSKVS 992  
 DB 944 VRDRLYASENHQELIKDKLILK:FEVLAPQVRYTY:EKHEMLPKSFIKLKSKVS 1002  
 QY 993 RFLRPYK 995  
 DB 1003 SFLRPYK 1009

## RESULT 14

US-10-239-663-58  
 : Sequence 58, Application US/10239663  
 : Publication No. US20030139572A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Agarwal, Pankaj  
 : APPLICANT: Murdoch, Paul R.  
 : APPLICANT: Rizvi, Safia, K.  
 : APPLICANT: Smith, Randall, F.  
 : APPLICANT: Xiang, Zhaoying  
 : APPLICANT: Kabnick, Karen  
 : TITLE OF INVENTION: NOVEL COMPOUNDS  
 : FILE REFERENCE: GPE0018  
 : CURRENT APPLICATION NUMBER: US/10/239, 663  
 : CURRENT FILING DATE: 2002-09-24  
 : PRIOR APPLICATION NUMBER: PCT/US01/09226  
 : PRIOR FILING DATE: 2001-03-22  
 : PRIOR APPLICATION NUMBER: 66/192,158  
 : PRIOR FILING DATE: 2000-03-24  
 : PRIOR APPLICATION NUMBER: 66/192,668  
 : PRIOR FILING DATE: 2000-03-27  
 : PRIOR APPLICATION NUMBER: 60/200,166  
 : PRIOR FILING DATE: 2000-04-27  
 : NUMBER OF SEQ ID NOS: 66  
 : SOFTWARE: FASTSEQ for Windows Version 3.0  
 : SEQ ID NO 58  
 : LENGTH: 964  
 : TYPE: PRT  
 : ORGANISM: Homo sapiens  
 US-10-239-663-58

Query Match 58.7%: Score 3290.5; DB 12: Length 964;  
 Best Local Similarity 59.7%: Fred. No. 2.5e-241;  
 Matches 580; Conservative 110; Mismatches 162; Indels 119; Gaps 14;

QY 10 GAA---AAVLLLLLLPPLLALAGAVPFGRGAG-----PQEDVDECAQGLDQCHADAL 61  
 DB 2 GAAVAMHLVLLAL-----GTRGLAGGSLPSSVVDDESSESTDCCHIDAI 49  
 QY 62 CQNTPTSYKSCCKPGYQGEHQCEDIDECNE-LNGCVHDCLN:PGNYRCTCFDGFMLA 120  
 DB 50 CQNTPKSYKCLCKPGYKSGKCEDIDECENDYNGCVHCCIN:PGNYRCTCFDGFMLA 109  
 QY 121 HDGHNCLVDECLENNCGQHTCVVMWSYECCEGEGFLSDNCHTCHRSSEGLSCNMK 180  
 DB 110 HDGHNCLVDECCQDNNGCQQT:CVAMWSYECQSGSGFLSDNCHTCHRSNENXNMK 169

QY 181 DECSHICKEAPRGSVACECRPGFELANQPDCL:TCNHNNGGCGHSCD:ADPECSCH 240  
 DB 170 DHCACI:CRFTPKGVACDCRPGFELANQPDCL:TCNHNNGGCGHSCD:EDPRTCCG 229  
 QY 241 PQYKMTDGRSCLEED:TVLEVTSSNTSVDGDKRVKRLMKETCAVNNGCCRTCD: 300  
 DB 230 QYALHSDGRTCI-----ETCAVNNGCCRTCD: 259  
 QY 301 STGVHSCSPGVFTJLDBKTCCKJDEQOTRNGSCDHPCKN:VGSFDCCKKGFLLTDEK 360  
 DB 260 ATGVHSCSPGVFTJQPDKTCCKJDEQOTRNGSCDHPCKN:VGSFDCCKKGFLLTDEK 319  
 QY 361 SCQVDGCSLDRTCHSCINHPGTACACNNGYTLGFTTHGDTNECS:INNGCCQCVN 420  
 DB 320 TQDDIDEGSFEXTCHICINSGS:QCCCHNGYTLGFTTHGDTNECS:INNGCCQCVN 379  
 QY 421 TVGSYECQCHPGYKLMNKDCVEV-KALLPVSVPVSLHCGSGGSDGCFRLC----- 474  
 DB 380 TKGSYECVCPGGRILHNNKDCVEV:KALLPVSVPVSLHCGSGGSDGCFRLC----- 439  
 QY 475 -----HSGHLS-----SDVTITRTSYTFKLNKCSGL 502  
 DB 440 FVPSNENSVYVSCVPGPQKALQKRNQTSGLGFCSDAP:TPIKQYARFKIRAKHL 499  
 QY 503 KNAELFPEGRLPALDEKHSVKESEFR-----VYNCTCSSGQVPGAGPSTPKE 552  
 DB 500 R-----FHSQARAKETARQPLDHCVTFTJLKCDSKXK--RRGRKSPSE 544  
 QY 553 X-FIVPEFELETNQEVLASCDLCTIVKTEKRLKARILKRAVH2EQFHLQSGMND 611  
 DB 545 VSHIAEFELTFKVEASDTCEDCLRKAEQSLQMA:KILKSKIGROCFYQVSGTEVE 604  
 QY 612 VAKKPRHTERAEQSGVGGGHAENQVSCARGTYDCAERKICLPGTFQNEGQVTC 671  
 DB 603 VQRPYKAL:EGQ-ACGAGQYVQDKCYACGGHTFSGELGQVCSCYETVDMGQSLC 663  
 QY 672 EPCPPRGNAGALKTEAMNMSECGSLCQPEYSADGAPACQCALGTFQPEAGRTSCPEPC 731  
 DB 664 TPCP---SSDGLCLGAGANVSECGSGCSGSPFSAAGCPQCCQCPVGTQPEAGRTSCPEPC 720  
 QY 732 GGGJATKQCATSPQDCE:TRVOCSPGCHRYN:TTTHAC:ACPVTQYPEFGKNNVCSPONT 791  
 DB 721 GGGJATKQCATSPQDCE:TRVOCSPGCHRYN:TTTHAC:ACPVTQYPEFGKNNVCSPONT 780  
 QY 792 STDPGOSTVTHCKKQKQCGGELGDTYGTIESPNYGCY:PAHNECTYHIAFPKRLIIV 840  
 DB 791 STDPGOSTVTHCKKQKQCGGELGDTYGTIESPNYGCY:PAHNECTYHIAFPKRLIIV 840  
 QY 852 PEIFLPIEDDCGCVLVMKRTSSSVTTEYETQTERP:AFTRSRSKSLMIOFKSNEGNSA 911  
 DB 841 PEIFLPIEDDCGCVLVMKRTSSSVTTEYETQTERP:AFTRSRSKSLMIOFKSNEGNSA 900  
 QY 912 KGFQVRYTYD 922  
 DB 901 KGFQVRYTYD 911

## RESULT 15

US-10-221-097-51  
 : Sequence 51, Application US/10221097  
 : Publication No. US20030144476A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Agarwal, Pankaj  
 : APPLICANT: Murdoch, Paul R.  
 : APPLICANT: Rizvi, Safia, K.  
 : APPLICANT: Smith, Randall, F.  
 : APPLICANT: Xiang, Zhaoying  
 : TITLE OF INVENTION: NOVEL COMPOUNDS  
 : FILE REFERENCE: GPE0016  
 : CURRENT APPLICATION NUMBER: US/10/221, 097  
 : CURRENT FILING DATE: 2002-09-26  
 : PRIOR APPLICATION NUMBER: PCT/US01/07143  
 : PRIOR FILING DATE: 2001-03-05

```

; PRIOR APPLICATION NUMBER: 60/187,107
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: 60/236,874
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/188,916
; PRIOR FILING DATE: 2000-03-13
; PRIOR APPLICATION NUMBER: 60/237,846
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 51
; LENGTH: 965
; TYPE: prt
; ORGANISM: Homo sapiens
; US-0-221-097-51

```

```

Query Match      58.7% Score 3290.5; DB 12; Length 965;
Best Local Similarity 59.7% Pred. No.2.5e-241;
Matches 580; Conservative 110; Mismatches 162; Indels 119; Gaps 14;

```

```

QY 10 GAA--MAVLLLLLPRLLLLAGAVPPGRRAAG-----PQEDVDECAQGLDQCHADAL 61
DB 2 GAAAVRMLCVLLAL-----GTRGLAGSGGLPSVADVDECEGDDCHIDAI 49
QY 62 CONTPTSYKCKPGYOGEGROCEDIDECGNE-LNGGCVHEDCLNIPGNYRCTCFEGMLA 123
DB 50 CQNTPKSYKCKPGYKSGKQCEDIDECENDYNGGCVHACINIPGNYRCTCFDGYLA 109
QY 121 HDGHNCLDVEDECLENNCGQHTCVVMVNGSYECCKCKESFFLSDNCHTCHHSEEGLSCKXK 183
DB 110 HDGHNCLDVEDEQDNGGCGQICVAMKSYECQCHSGFFLSDNCHTCHHSEEGLSCKXK 169
QY 181 DHGSHICKAPRGVACECRPGFEELAKNGQDCITLCHNGNGGCGHSCDTADGPBSCSH 240
DB 170 DHGCAHICRETPKGGVACDCRPGFDIAQNKDCTLTCHVNGNGGCHSCEDTDTGPTCCH 229
QY 241 POYKMHTRGSCLEHEDTVLEVESNTTSVVDGDKVAKRLLMETCAVNNGGCDRTCKDT 300
DB 230 QKVALHSDORTCI-----ETCAVNNGGCDRTCKDT 259
QY 361 STGVAGSCPVGFTLLDQKTKCDIDECQTRNGGCDHFCNKIVGSFDGCKKFKLLTDEK 360
DB 260 ATGVAGSCPVGFTLLDQKTKCDIDECQTRNGGCDHFCNKIVGSFDGCKKFKLLTDER 319
QY 361 SCQDVDECSLERTCHSCINHPGTFACACNMGYTLXGTHCGDTHNCSINNGGQCVCKN 420
DB 320 TCQDIDECSEFERTCHICINSPGSFQCLCHRGYLLXGTHCGDVECSMSKSGSCDQGCYN 379
QY 421 TVGSYECQCHPGYKLMWKKDCVEY-KGLLFVSPPVSLHCGKSGGGGCGFLRC----- 474
DB 380 TKGSYECVCPGRRLHNMKDCVETGKCLSRKTSPPRAQLSCSKAGVESCFLSCPAHTL 435
QY 475 -----HSGIHLSS--SDVTITRTSVTFKLENGKCSJ 502
DB 440 FVPDSENSVYLSGVPGFGKALQKRNSTSSGLSPSCSDAPTPIKOKARPKIRDAKCHL 499
QY 503 KNAELPREGRLALPEKKSVSYSER-----YVNLTCSSGKQVPGAAPRSTPKE 552
DB 500 R-----PHSQARAKETARQFLDHCHVTFTLKCDSKK--RRGRKSSPKE 544
QY 553 M-FIVEFELETNQKEVASCGLSCIVKTEKRLKAIKTRKAVHREQFHLQSGNNLD 611
DB 545 VSHITAEFEIETKHEASDTCGADCLRKRAEOSCALIKTKRSGRQGFVQVSGTEYE 604
QY 612 VAKKPPRTSEROAESCGVQGHAEQNCVSCRAGTYVDGARERCICPNNGTFQNEEGQMT 672
DB 605 VAORPAKALEGGG-ACGAGQVLDQSKCYACPGTHFGSELGQCVSCMPGTQDMWEGQJSC 663
QY 672 EECPRRNSGALKTPLEANMSECGGLQSGEVSADGFAPQCLALGTFQPAAGRTSGEPC 731
DB 664 TPCP---SSDGLGLPAGANVSECGGCGSPGFFSADGFRPCQACPVGTYQPEPGRGCGPC 720
QY 732 GGGLATKHGATSPQDCETRVGCSPGHFYNTTHRCICRPVGTYPQPEFGKNNVSCPANT 791

```

```

DB 721 GGGLATKHGATSPQDCETRVGCSPGHFYNTTHRCICRPVGTYPQPEFGQNHCTCPGNT 780
QY 792 TTPDQSTNITQCKXRRCCGELGDFGYIESPNYPGNYPANTECTWTINPPKRRLIIV 851
DB 781 STDPDSTNIVTHCKNQHCGGELGDTGYIESPNYPGNYPANAECVWHLAPPKRRLIIV 840
QY 852 PEFLPTEEDCGDYLVMKRTSSSNVTTYETQTYERPIAFTSRKSLMICKFKNESGSA 911
DB 841 PEFLPTEEDCGDYLVMKRSAPTSITTYETQTYERPIAFTSRKSLMICKFKNESGSA 900
QY 912 RGFQVPTVTTD 922
DB 901 KGFQVPTVTTD 911

```

```

Search completed: October 22, 2003, 11:40:28
Job time : 97 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: October 22, 2003, 11:25:58 ; Search time 31 Seconds

(without alignments)  
1363.502 Million cell updates/sec

Title: US-09-747-371-2

Perfect score: 5605

Sequence: 1 MGAVGRNRRGAAWAVLLLL.....PRFIRLRVNSFLRPYX 999

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCITUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match length	DB ID	Description
1	692	12.3	956 2	US-08-897-443-3 Sequence 3, Appl 1
2	682.5	12.2	638 2	US-08-897-443-1 Sequence 1, Appl 1
3	624.5	11.1	1394 6	517197-30 Patent No. 5,771,97
4	571	10.2	1253 3	US-08-479-722B-4 Sequence 4, Appl 1
5	553.5	9.9	1251 5	PCT-US95-02251-3 Sequence 3, Appl 1
6	553.5	9.9	1252 1	US-08-199-780-3 Sequence 3, Appl 1
7	553.5	9.9	1252 2	US-08-316-650-3 Sequence 3, Appl 1
8	552	9.8	1833 5	US-08-479-722B-2 Sequence 2, Appl 1
9	552	9.8	1833 5	PCT-US95-02251-18 Sequence 18, Appl 1
10	526.5	9.4	2321 4	US-09-230-652-2 Sequence 2, Appl 1
11	523.5	9.2	2703 1	US-08-185-432-19 Sequence 19, Appl 1
12	523.5	9.2	2703 4	US-08-899-232-4 Sequence 4, Appl 1
13	501.5	8.9	2523 1	US-08-185-432-18 Sequence 18, Appl 1
14	501.5	8.9	2523 4	US-08-899-232-3 Sequence 3, Appl 1
15	491.5	8.8	2556 1	US-08-083-590A-20 Sequence 20, Appl 1
16	491.5	8.8	2556 3	US-08-532-384-20 Sequence 20, Appl 1
17	485	8.7	2471 1	US-08-185-432-16 Sequence 16, Appl 1
18	485	8.7	2471 1	US-08-083-590A-19 Sequence 19, Appl 1
19	485	8.7	2471 4	US-08-532-384-19 Sequence 19, Appl 1
20	485	8.7	2471 4	US-08-899-232-1 Sequence 1, Appl 1
21	483	8.6	2556 1	US-08-185-432-17 Sequence 17, Appl 1
22	483	8.6	2556 4	US-08-899-232-2 Sequence 2, Appl 1
23	482.5	8.6	1964 4	US-09-467-997-1 Sequence 1, Appl 1
24	431	7.7	1055 3	US-09-214-278-2 Sequence 2, Appl 1
25	430	7.7	1065 2	US-08-400-159-8 Sequence 8, Appl 1
26	429	7.7	1212 3	US-09-214-278-3 Sequence 3, Appl 1
27	429	7.7	1257 3	US-08-611-729A-8 Sequence 8, Appl 1

28	427	7.6	1238 3	US-09-214-278-5 Sequence 5, Appl 1
29	426.5	7.6	448 2	US-08-984-072-1 Sequence 1, Appl 1
30	426.5	7.6	448 2	US-09-212-168-1 Sequence 1, Appl 1
31	424	7.6	1193 2	US-08-400-159-10 Sequence 10, Appl 1
32	424	7.6	1193 3	US-08-611-729A-10 Sequence 10, Appl 1
33	422.5	7.5	1218 2	US-08-400-159-6 Sequence 6, Appl 1
34	422.5	7.5	1218 3	US-08-611-729A-6 Sequence 6, Appl 1
35	421.5	7.5	1218 3	US-08-882-046-7 Sequence 7, Appl 1
36	421.5	7.5	1236 4	US-09-068-740A-6 Sequence 6, Appl 1
37	421.5	7.5	1187 4	US-09-068-740A-7 Sequence 7, Appl 1
38	421.5	7.5	1208 4	US-09-199-865-1 Sequence 1, Appl 1
39	421.5	7.5	1218 3	US-08-882-046-2 Sequence 2, Appl 1
40	421.5	7.5	1218 4	US-09-068-740A-11 Sequence 11, Appl 1
41	419.5	7.5	1218 3	US-09-214-278-7 Sequence 7, Appl 1
42	412	7.4	1219 3	US-08-882-046-5 Sequence 5, Appl 1
43	412	7.4	3623 2	US-09-141-461-2 Sequence 2, Appl 1
44	410	7.3	443 4	US-08-833-963C-2 Sequence 2, Appl 1
45	410	7.3	443 3	US-08-980-514-1 Sequence 1, Appl 1

## ALIGNMENTS

RESULT 1  
US-08-897-443-3  
Sequence 3, Application US/08897443  
Patent No. 5,991,263  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Lal, Preeti  
APPLICANT: Corley, Neil C.  
APPLICANT: Shah, Puri  
APPLICANT: Kaser, Matthew  
TITLE OR INVENTION: HUMAN MACRILIN-3  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/897,443  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy C.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0348 US  
TELEPHONE: 415-855-0555  
TELECOMMUNICATION INFORMATION:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 956 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: Genbank  
CLONE: 2072792  
US-08-897-443-3

Query Match 12.3% Score 692, DB 2, Length 956,  
Best Local Similarity 29.8% Pred. No. 1.7e-4.

Matches 159; Conservative 72; Mismatches 224; Indels 78; Gaps 14;

QY 62 CONPTSYKSCCKPGY--QGEGRQCEDIDECNELNGCVCHDCLNIPANRYCTGDFGML 119  
 DB 253 CLNTGSIYCKCKQGYILSTQKCRIDLCATE-DHCEGLCVNMGSPFVQCYSYTL 311  
 QY 120 AHDGNCIDVDECLNNGCCQHTCVNMGSEYECCKEGFSLSDNCHTCHRSBGLSCYN 179  
 DB 312 AEDGKCTAMDYCASENHGCHECNMAESSYLCHCEGFAINSKCKTCKID---YCAS 367  
 QY 180 KDHGSHCKEAPRGSVACGRPGFELAKNORDC--ILTCHNGCCQHSODDTADSPFC 237  
 DB 368 SNHGQHECVNA-QTSALCRCLKGFMILPDKKTRINVCALNKGCEHECVNTEBHYC 426  
 QY 238 SCHPOYKMTDGRSCJE-----REDTLEVTESNTTSVVDG--DKVYRRLME 284  
 DB 427 RCRGGYNLDRPKGKTSRVDHCAQDHGCEQCLNTEEFVCCSGSEGLINDLKTCSRAC 456  
 QY 285 TCAYNNGGCDRTCKTSTGVHSCFVFTLQDGTCKCJDECCTRNGGCDHFCANVGS 344  
 DB 487 YCJLSNHGCEYSCVNTDKSFACQCFEGHVLRSDDKTKAKLDSGALGDHGEHSCVSSDS 546  
 QY 345 PDGCKGKFXLLTEPKSCQVDEC-SLDRDCHSCINHPGTFACACRGYTL-YGFTHCG 402  
 DB 547 FVCCQFEGYILRDGKTKRKADVCODVHGEHLCAVSGSEYVCKLBFRLAEDGRCR 606  
 QY 403 DTNECSINNCGCQGVCVNTVSGSYECQCHPGYKLMNKKDC-----444  
 DB 607 RKNVCKSTQHGEHNCVANNNGNSYLCRCSEFVLAEDGHCRCRTEGPIDVFFV-DGSKSL 666  
 QY 445 -----VKGLLFT-SVSPRSLHCGSGSGGDFLCHGCHILSSVITPRISVT 492  
 DB 667 GEENFETVGHFVTGLIDSLAVSPKA-----ARVGLLOYSTQVTEFT 728  
 QY 493 FK---LNEGKSLKNAELPFGALRPALPEKHSYKSEFRVYVNTCSGKQVP 541  
 DB 709 LRGFSSAKEMKAVYHMKYMGKSGMTGLALGH-MERSFTQVBGARPRSTQVP 760

RESULT 2  
 US-08-897-443-1  
 Sequence 1, Application US/08897443  
 Patent No. 5981263  
 GENERAL INFORMATION:  
 APPLICANT: Hillman, Jennifer J.  
 APPLICANT: Lal, Preeti  
 APPLICANT: Corley, Neil C.  
 APPLICANT: Shah, Purvi  
 APPLICANT: Kaser, Mathew  
 TITLE OF INVENTION: HUMAN MATRILIN-3  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 STREET: 3174 Porter Drive  
 CITY: Palo Alto  
 STATE: CA  
 COUNTRY: USA  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDICAL TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSO for Windows Version 2.3  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/897,443  
 FILING DATE: Filed Herewith  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Billings, Lucy C.  
 REGISTRATION NUMBER: 36,749  
 REFERENCE/POCKET NUMBER: PF-0348 US

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-855-0555  
 TELEFAX: 415-845-4166  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 638 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: JTRSNCT2  
 CLONE: 681719  
 US-08-897-443-1

Query Match 12.2%; Score 682.5; DB 2; Length 638;  
 Best Local Similarity 33.0%; Pred. No. 5e-41;  
 Matches 139; Conservative 60; Mismatches 163; Indels 59; Gaps 11;

QY 62 CONPTSYKSCCKPGY--QGEGRQCEDIDECNELNGCVCHDCLNIPANRYCTGDFGML 119  
 DB 253 CLNTGSIYCKCKQGYILNSTDTCRIDLCAE-DHCEGLCVNMGSPFVQCYSYTL 311  
 QY 120 AHDGNCIDVDECLNNGCCQHTCVNMGSEYECCKEGFSLSDNCHTCHRSBGLSCYN 179  
 DB 312 AEDGKCTAMDYCASENHGCHECNMAESSYLCHCEGFAINSKCKTCKID---YCAS 367  
 QY 180 KDHGSHCKEAPRGSVACGRPGFELAKNORDC--ILTCHNGCCQHSODDTADSPFC 237  
 DB 368 SNHGQHECVNTD-DYSCHCLKGFTLINPDKKTRINVCALNKGCEHECVNTEBHYC 426  
 QY 238 SCHPOYKMTDGRSCJLEREDTLEVTESNTTSVVDGRVYRRLMEFCVANNNGGCDRT 297  
 DB 427 RCHRGYLDPNKTKTSR-----VDHCAQDHGCEQLC 458  
 QY 298 KDTSTGVHSCFVFTLQDGTCKCJDECCTRNGGCDHFCANVGSFDCGCKKFKLLT 357  
 DB 459 LNTDSFVQCSEGLINDLKTCSRVDYCLSDHGEYSCVNMGRSFAQCQFESHVLS 518  
 QY 358 DEKSCQDVDEGL-DRTDHCSCINHPGTFACACNRYTL-YGFTHCDTNECSINNQC 414  
 DB 519 DGTCAKLDSGALGDHGEHSCVSSDSFVCCFEGYILREDGK-CRRKVCQAIDHQC 577  
 QY 415 QQVCVNTVSGSYECQCHPGYKLMNKKDC-----EVVGLPITSVPVS 458  
 DB 578 EHCVNSDDSYTCEGLGRLAEDGRCKRRYSANOPTMAANTFVLNG-IPSSANAQD 636  
 QY 459 L 459  
 DB 637 L 637

RESULT 3  
 517197-30  
 Patent No. 517197  
 APPLICANT: KANZAKI, TETSUO; OJOFSSON, ANDERS; MOREN, ANITA;  
 WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSEN, WELSH,  
 LEVA, HEIDIN, CARL; HENRIK  
 TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING  
 HUMAN TRANSFORMING GROWTH FACTOR-BETA1-BINDING PROTEIN  
 NUMBER OF SEQUENCES: 53  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/487,343  
 FILING DATE: 27-FEB-1990  
 SEQ ID NO: 30  
 LENGTH: 1394  
 517197-30

Query Match 11.3%; Score 624.5; DB 5; Length 1394;  
 Best Local Similarity 24.9%; Pred. No. 1.9e-36;  
 Matches 240; Conservative 94; Mismatches 334; Indels 297; Gaps 48;

QY 45 DVECAQGDGCGIADALCONPTSYKSCCKPGYQ--GGEGRQCEDIDECNELNGCVCHD- 101



Db 546 EINECTVNPDIQGA-GHCINLPVRYTCIYEGYRSEQKCKVCVIDEC-TQVCHLSCGR 603  
 QY 102 CLINPGNYRCTCFDGFMLAHGHNCLVDECELENNKGCCHTCVNMVSGYEC-CKEGRFEL 163  
 Db 604 CENTGSGFLC:CPAGFPASBEGTNCIDVDECBRLPVCGEGCVNIVGAFREYCDGSGRM 663  
 QY 161 SDNQHTCIHRSEEGSLCMKHCSCSHICKAPRGSVACECPGFEIARQPCILTCNHG 220  
 Db 664 TOR-----GRCEIDECINPSTCPBQCVNSP-GSYQC-----VPCREG 701  
 QY 221 NGCQHSODTADGPECSCHPOYKMHDTGRSCLEPEDIVLEVTSNTTSVVDGDKRVKR 280  
 Db 702 FPGMNGQCLDVE-----CLF----- 717  
 QY 281 LLMETCAVNNGGCDRTCKDTSTGVHSCQVGFLLDAGTKCDIDECOTRNGCDHPCKN 340  
 Db 718 --PNVCA--NG--DCSNLESGTWSCHKGTTRPDHKGCDIDECQGLCVNGQCKN 769  
 QY 341 IVGSPDCGCKKGFLLTDEKSGQDVDECSLDRTCDH-SCINHPGTFAACAKRQYTLVGF- 398  
 Db 770 TEGSFRCTCGQGYOLSAKQDCEIDECQHRLLCAHQGQRTTEGSPQCVCQGYRASGLG 829  
 QY 399 THGSDTNECSINNCGCQ-VCVNTVSGTECQCHPGYKLMNK-----KQCVKGLPFS 452  
 Db 830 DHCEDINECLEBXSVCORQDCINTAGSYDCTCPDGFOLDNMTCODINEC-EHPGL--- 894  
 QY 453 VSPRYSLHQKSGS---GGDGGF-LHCHSG:HLSSVTTIR:SVTKLNEKCSLKNALF 508  
 Db 895 -----CGPGGECINTGSHFVCVCGQGFSTISADGRT-----CDI 918  
 QY 509 PEGLRPALPEKH---SSVKEFPYVNLTCSSGKQVP-----GAPGRP- 547  
 Db 919 DECVNNTVCDSHGFNDTAGSPRL---CYGQFQFPQGGQCVVNECELLSGVGEAFC 975  
 QY 548 S-CPKEMFIV-----EFLETNQKEVTASCDLSIVKTEKPLKKA:RT:RKA----- 595  
 Db 976 ENVEGSLFCVACADENQEVSPMTGQCRSRTSLDLDVAVQCPKEEKKCYVNLNDASLCNV 1035  
 QY 596 ----VHRBGFHLQLS---SMNLDAKKPRTSERGAESGVCVGGH-----AEN- 636  
 Db 1035 LAPNVTKQSCCTSGAKMGWDCNCEIFPCFVLGTAEFTEKCPKKGFPVAGESSSEAGGENY 1095  
 QY 637 ----QCV-----SCRAGTYVDGARRECI-----JCPNGTF 662  
 Db 1096 KMADECLLFGOEICKNGFLNTRPGVECYCKQGTYYDPVKLOCFPMDECCQPSIDCQC 1155  
 QY 663 QNEGQMTG-----EPCPPRGNS-----G 681  
 Db 1156 VNTGGSVNGFCTHPVWLDASEKRCIRPAESNEQ:EBTDVYCDLWEHLSDVEVCSRPVWG 1215  
 QY 682 ALKT-----PEAMNMSGGGLGQPGEYSADGAFPOQLCALSTFCEAGRTSGF--- 729  
 Db 1216 KQTTYTECCCLYGEAMGX-QC-ALC-PLKCDSDVQCLNIPVGTGRCP-YGRDALVDPE 1272  
 QY 733 ---PCGGGLATGAGATSFQDCETRVQCSPGHFVYTTTHRC:RCFVG-----TYCEP 779  
 Db 1272 QYTFEADPYFIQDRPLANSFELQAE-EC-G:LNGENGRCVRVDEGTCDCLDGYHLDT 1328  
 QY 780 GKNKCVSCPGNTTTTJPDGSTN-ITQCKNRCGGELGDF-----TGYES-PNY-PGNY 829  
 Db 1329 AKMTCFD-----VNECEDELNRMSLCKNAKCIINTGYSKCLCLPGVVPBPKNYCTPANT 1383  
 QY 830 PANTJE 834  
 Db 1384 A:ANLE 1388

APPLICANT: Yin, Wushan  
 TITLE OF INVENTION: LATENT TGF- $\beta$  BINDING PROTEIN (LTBP)  
 TITLE OF INVENTION: GENES, COMPOSITIONS AND METHODS  
 NUMBER OF SEQUENCES: 13  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Williams, Morgan & Amerson  
 STREET: 7676 Hallmont, Suite 250  
 CITY: Houston  
 STATE: Texas  
 COUNTRY: USA  
 ZIP: 77040  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/479,722B  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US PCT/US95/02251  
 FILING DATE: 21-FEB-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 09/316,650  
 FILING DATE: 30-SEP-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/199,780  
 FILING DATE: 18-FEB-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Fussey, Shelley P.W.  
 REGISTRATION NUMBER: 39,458  
 REFERENCE/DOCKET NUMBER: 4100.000500/FUS  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (713) 934-7200  
 TELEFAX: (713) 934-7011  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1253 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-479-722B-4  
 Query Match: 10.2%; Score 571; DB 3; Length 1253;  
 Best Local Similarity: 24.5%; Pred. No. 1, 1e-12;  
 Matches: 218; Conservative: 83; Mismatches: 296; Indels: 292; Gaps: 46;  
 QY 32 VPPQGRG---AAGQEDVDECAQGLDDCHADALQNTPTSYKSCPKRYG--EGRQCED 86  
 Db 539 LPEPSASVELAPQVETDECRNLNQC-DHQCVQPPSDYSCHNAGYRSHPCRYCVD 596  
 QY 67 IDECGNLNGCVCHDC:NTGNVTRCTCFDGFML--AHGHNCLDVEDC---LENNGQC 140  
 Db 597 VNEEAPPCPGKQKICNTGSYNCHNRGRYRVAGGASCVLLNECAKPHLCGGDG-- 654  
 QY 141 HTCVNMGVSECCCKEGFFLSDQHTCIHRSEGLSCMN:DHGSHICKZAPRSVAC-E 199  
 Db 655 -FCINFGHYKCNCPGYRLKASRPICEDIDE---CRDSTGDCGCKPK--GSFQCIA 709  
 QY 200 CRPFEEAKNQRDICILTCNHNQGCQHSQCDTADGPECS---CHPGYKMHDTGRSCLER 255  
 Db 710 CQPGYR-----SGGGGACR-----DVNECSGTPCSPGM----- 738  
 QY 256 EDYVLEVTSNTTSVVDGDKVKKRLME:CAVNNGSCDRTCKDTSTGVHSCQVGFLLQ 315  
 Db 739 -----CEVLPSSYRCTCAQGIATR 757  
 QY 316 LDGKCKDIDECQTRNGSCDHPCKNIVGSPDCGCKKGFLLTDEKSGQDVDECSLDRTG- 374  
 Db 758 TGRSLCIDVDECEGKVCQDQIGCTNTGSGTCCCLSYHL:SRPSRREDIDECDFPACI 817  
 QY 375 DHSCTNHGTFAACAKRNGYTLVGFTHC-GDTNECSINNQC-QQVCVNTVGSYECQCHPG 432

RESULT 4  
 US-08-479-722B-4  
 Sequence 4, Application US/08479722B  
 Patent No. 6074840  
 GENERAL INFORMATION:  
 APPLICANT: Bonadio, Jeffrey

```

Db      818 GDDCINTNGSYACCLPGLRVLGGRKCKKDDIDECSDPGLCLPACNELQSSYVCVDEG 877
Qy      433 YKLHNKKKDCVEVKGLLPTSVSPRVLCHGKSGGDDGFLRCHSGIHLSSDVTIRTSVT 492
Db      878 FTLDQDHGCEVE-----QPHHKK-----CYLNPDCVFPDS---VLATNVT 919
Qy      493 FVLNKGCSL-----KNMELPEEGLRPALPEKHSVKESFHYVNLTCSSGQVPAQAPRP 547
Db      919 --QDECCSLGAGMGDHCEIYPCPYSS-AEFHSLVFGKR-----LHSGQO----- 962
Qy      548 SPPKEMITVERELETNCKEVTASCDLSQIVKRTKRLRKALPLTKRAVHECHLQLDSG 607
Db      963 -----HCEL-CI-----PAHRCIDECILFG 981
Qy      608 XALDVAKRPRTSERQAESGCGVGHAEKQVS-----CRAGTYDGAERPCI----- 655
Db      982 -----AEIC-----KEGKCVNTQPGVECYCKQGFYIDGMLLECVUDE 1019
Qy      656 ----LCPNGTFQNEGGQMTCEPCP---RPGNSGALKTPPEANXSE---CGG-CQPGE 702
Db      1020 CLDESNGNVCENRGYRCACTPPAEYSPAQACL-IPERWSPGRDVKACAGASE--E 1076
Qy      703 VSADGAPQCALGTF-----QPEAGRTSCFPC-----GGGLATKIQGATSPQDC--- 748
Db      1077 RYACVGMWAPAL-TPDDCCCRPRLG-TQCRPCPPRGTSQCTSSSENSTFDTSP, 1134
Qy      749 -----ETRVQCSPGHFYNTTTRCI-----RCPVGTIYQPFQKNCVSC 787
Db      1135 LIGKSPRDESSSESDDEC-----RCVSGRCVPRPGAVCECP-GGQLDASRARCV-- 1185
Qy      788 PGNITTPFGSTNITQ---CKNRRCGELGDF-----TGVISS-PNYP 826
Db      1186 -----DIDECRELNRGLLCKSERCVNTSGSFRCVAKGFTSRPHGP 1228

RESULT 5
PCT-US95-02251-3
Sequence 3, Application PC/TUS9502251
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STIMULATING BONE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: Patent Release #1.0, Version
SOFTWARE: #1.38
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/02251
FILING DATE: CONCURRENTLY HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/316,650
FILING DATE: 30-SEP-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Parker, David L.
REGISTRATION NUMBER: 32,165
REFERENCE/DOCKET NUMBER: UMIC009P-
TELECOMMUNICATION INFORMATION:

```

```

TELEPHONE: (512) 418-3000
TELEFAX: (713) 799-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1251 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-02251-3

Query March 9, 98; Score 553.5; DB 5; Length 1251;
Best Local Similarity 24.63; Pred. No. 21e-31;
Matches 218; Conservative 82; Mismatches 300; Indels 285; Gaps 47;

Qy      32 VPPGRGR---AAGPGEVDDECAQZJDDCAGALACONTPTSYKSCSKPYQG--EGROCED 96
Db      537 LPPSRSAVBLAPQVTEDEBCLNQNIC-GHQCVPGPBDYSCHNAGYRSHPHRYCYD 595
Qy      27 IEQGNELNGCCVHDCCUNIPGNYRCTCFDGFML--AHQHNCLVDDEC---LENNGGCQ 140
Db      596 VNECEAEPPGPGKGIOMNTGGSYNCHNGRYLHYAGAGRSQVDLNECKKPHLCGGG-- 653
Qy      141 HTGVNMGSECCCKEGFELSONHTCIHSEEGSCNKHGQCSHTCKEAPRGSYAC-E 199
Db      654 -FCINFGHYKCYGYRLKASRPICEDIDE--CRPSTCPGCKCNKP-GSFKCA 702
Qy      200 CRPFELAKNRDCLITCNHNGCCQHSQCDTADGPGSCHEQYXNHTDGRSCLREEDTV 259
Db      709 CQGYR-----SQGGAGR-----DVNEGS-----EGTFC----- 733
Qy      260 JEVTSNTTSVVDGAKRYRRLMETCAVNNNGCCTRTCKDTGVNCSGPFVFTLDLDR 319
Db      734 -----SPWCKE---LPGSYRCTAQGIRTRGR, 760
Qy      320 TGKDIDECOTRNGGCDHPCKNIVGSPDCGKXGFKLITDEKSCQVDECSLDRTC-DHSC 378
Db      761 SCIDVDDCEGKXVCGGICTNTPGSGQCLSGYHLSRPSKCEDIDEIDPAACTGGDC 820
Qy      379 ZHNGTFACACNGYLYGFTHC-GEINBGS-INNGC-QQVAVNTVGSYECCQHPYKXH 436
Db      821 INTNGSYRCPLGRLHYGRKCKKDIDCSQDPLCLPHACENLQGSYVCVDEGFTLG 880
Qy      437 MNKDCVEVKGLLPTSVSPRVLHCKSKSGGCGCFRCHSGIHLSSDVTITRSYTFKXN 496
Db      881 QDHGCEVE-----QPHHKK-----CYLNPDCVFPDS---VLATNVT--QQ 919
Qy      497 EGKCSL-----KNMELPEGLPALPEKHSVKESFHYVNLTCSSGQVPAQAPRSTPK 551
Db      920 ECCCCLGAGKDHCEIYPCPYSS-AEFHSLVFDGR-----JHSGQO----- 961
Qy      552 EMETIVEELETNCKEVTASCDLSQIVKRTKRLRKALPLTKRAVHECHLQLDSGLD 611
Db      962 -----HCEL-CI-----PAHRCIDECILFG----- 980
Qy      612 VAKKPRITSERQAESGCGVGHAEKQVS-----CRAGTYDGAERRC----- 655
Db      981 -----AEIC-----KEGKCVNTQPGVECYCKQGFYIDGMLLECVUDE 1022
Qy      656 -LCPNGTFQNEGGQMTCEPCP---RPGNSGALKTPPEANXSE---CGGLCPGFSYAD 706
Db      1023 SNCRNVCEN-TRALPCACPPAEYSPAQACL-IPERWSPGRDVKACAGASE-ERTAC 1076
Qy      707 GFAPQCALGTF-----QPEAGRTSCFPC-----GGGLATKIQGATSPQDC----- 748
Db      1079 VMGPWAPAL-TPDDCCCRPRLG-TQCRPCPPRGTSQCTSSSENSTFDTSP, 1136
Qy      749 -----ETRVQCSPGHFYNTTTRCI-RCPVGTIYQPFQKNCVSCQGN-----T 791
Db      1137 SPDEDSSESDDEC-----RCVSGRCVPRPGAVCEP-GGQLDASRARCV 1191
Qy      792 TIDFGSTNITQ---CKNRRCGELGDF-----TGVISS-PNYP 826

```

Db 1182 CVDIDECRELNGRLCKSERCVNTSGSFRVCVAKGFTSRPHGP 1226

RESULT 6  
US-08-199-780-3

Sequence 3, Application US/08199780  
Patent No. 5763416

GENERAL INFORMATION:

APPLICANT: Bonadio, Jeffrey

APPLICANT: Goldstein, Steven A.

TITLE OF INVENTION: Gene Transfer Into Bone Cells

TITLE OF INVENTION: And Tissues

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: TX

COUNTRY: USA

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30B

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/199,780

FILING DATE: 18-FEB-1994

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Parker, David

REGISTRATION NUMBER: 32,165

REFERENCE/DOCKET NUMBER: UMIC:002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 320-7220

TELEFAX: (512) 474-7577

INFORMATION FOR SEC. ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1252 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-199-780-3

Query Match 9.9%; Score 553.5; DB 1; Length 1252.

Best Local Similarity 24.6%; Pred. No. 2.1e-31;

Matches 218; Conservative 82; Mismatches 300; Indels 285; Gaps 47;

Db 32 VPRGRGR---AAGPQDVDECAAGLDDCHADALCONTPSYKSCCKPGYQ--FGQCED 86  
Db 538 LPPSRSAVEIAPQVETEDCRLNNTIC-GHGCQVPGPSDVSCHMAGYRSHPHRYCVS 596  
QY 87 IDECGNELNGGCVHDLNIPGNRYCTCFDGMFL-AHDGNCIDVDEC---LENNGCQ 140  
Db 597 VNECEAPGCPGKGIOMTGGSYNCHNNGYRLHVGAGRSCTDLNECKAPHLGCGG-- 654  
QY 141 HTCVNMGSEYECCKGFFLSNQHTCIHRSEBGLSCMKKQCSHICKEAFRGSVAC-E 199  
Db 655 -FCINFGHYKCMCYPGYRLKASRPICEIDIE---CRDPSTCPDGCKENKP-GSFKCIA 709  
QY 200 CRPGFELAKKQRCILITCNHNGSCQHSQCDTADGPEGCHPYKNTHTGRLSHEDIV 259  
Db 710 QDPGR---SOGGACR---DVNECS---EGTFC----- 734  
QY 260 LEVTSNTSVVDGKRVKRLIMETCAVANGCDRTCTDTSFGVCSQPVFTIQDGR 319  
Db 735 -----SPGCKE-----LPSYVCTCAQGRITRTGL 761  
QY 320 TCKDIECCGRNGSCCHFCNNIVGSPDCCGCKGFKLLTDEKSCQDVDEGSLDRFG-DHSC 378  
Db 762 SCIDVDDECEAGKVCQDGICTNTPGSPCCQLSGYH-SRDRSRCEIDIEDCPAACGCGD 821  
QY 379 INRGTACACNNGYTLTGFTHC-GDINBESINNQC-QQVCNVTGVSIECCGHPYKHL 436

Db 822 INTNGSRCLCPGRLHVGGRKCKKIDECSDQPGCLPHACENLGGSYVCCDEGFTLT 881

QY 437 WKKDCEYKGLPTSVSPVSLHCKSGGGGCGFCRCHSGHLSSDVTTRTSVTFKLN 496

Db 882 CDQHGEVE-----QPHHKE-----CYLNPDDTDFCDS--VLATNVT--QQ 920

QY 497 EKGCSL-----KNAELFPEGLRALDEKSSVSEFRVYNLTSSGKQVGAAGRSTPR 551

Db 921 ECCCSGAGMGDHCEYPCPVYSS-AEFSHLPDGR-----LHSGQ----- 962

QY 552 EXFIVEFELETNQKEVTASCDLSIVKTRKALKRIRLKAHVRQPH-OLSGMNLJD 611

Db 963 -----HCEI-CI-----PAHRIDECLFG----- 981

QY 612 VAKKPRTSERQAESGCVSGHAENQCVS-----CRAGTYDGRERCI----- 655

Db 982 -----AEIC-----KEGKCVNSQPGVEGCKGPFYDGRLECVDVDECLDE 1023

QY 656 -LCFNGTFQNEBQKTCPCP-----RPGNGALKTPEAMMSE-----CGGLCPGEXSAD 706

Db 1024 SNRNGVCN-TWRLPACTPAEYSPAQOCL-IPRWSLPQDYKAGASE--ERTAC 1079

QY 707 GFAPCOLCAUGTF-----QFEAGRTSCFPC-----GGG-ATRGAGATSPDC----- 748

Db 1080 VMGWKAPAL-TRDCCCRQPRUG-TQCRCPFRGRTGSGCPRTGSEBNSWDTSPLLGX 1137

QY 749 -----TRVQCSFGHRYNTTHRCIRCPVGTQPEFGKNCVSCPN-----T 791

Db 1138 SPREDSESEDEEC-----RCVSGPCVPRGAWC-ECPPGFLDASRAR 1182

QY 792 TTFPGSNTITQ-----CKNRGSGELGCF-----TYIES-PNYP 826

Db 1183 CVDIDECRELNGRLCKSERCVNTSGSFRVCVAKGFTSRPHGP 1227

RESULT 7  
US-08-316-650-3

Sequence 3, Application US/08316650  
Patent No. 5942496

GENERAL INFORMATION:

APPLICANT: Bonadio, Jeffrey

APPLICANT: Roessler, Blake J.

APPLICANT: Goldstein, Steven A.

TITLE OF INVENTION: METHODS AND COMPOSITIONS

TITLE OF INVENTION: FOR STIMULATING BONE CELLS

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: USA

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/316,650

FILING DATE: 30-SEP-1994

CLASSIFICATION: 514

PRIOR APPLICATION NUMBER: US 38/199,780

FILING DATE: 30-SEP-1994

ATTORNEY/AGENT INFORMATION:

NAME: Parker, David L.

REGISTRATION NUMBER: 32,165

REFERENCE/DOCKET NUMBER: UMIC:008

TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000

Db 1344 CG-----EETSPSGNECVDNCELMMAVCGDALCENVEGSPJC 1182  
 Cy 588 A:RTLRKAVHREDFHL--QLSGMNLDAVAKKXPRTSERGAES-----CG-- 628  
 Db 1383 LCASLAEYDAEBCHCRFRVAG---AQRIPVETEDQAPSLIRMECYSHNGGPPCSQ 1438  
 Cy 629 VGGHAENOCVSCAGTYDGAERECILCPNGTFQNEEGMTCEPCPR-----DG 673  
 Db 1439 LGQNSTQAECCCTGGARWGKA-----CAPGSEDSVEFSQICPS 1477  
 Cy 679 NSGALKTPBEA-----NMSEC---GGICQGEYSADGFAPCQICALGTGPEAGRT 726  
 Db 1478 GGGYLPVEGAMTFGGTMYTDADECVLGSPALCQGRCS--NIVPGYICL----- 1524  
 Cy 727 SCFPGGGGLATKRGATSFQDCETRVCCSPGHFNKTTTHRC-----IRCPVCTVQSE 778  
 Db 1525 -----CNPQYHYDASRRKQCHNECQDLACENGECVNO 1557  
 Cy 779 FGKXNCVSCPGNTTIDPDGNTITQCKARRCGELGDF-----TGYSER 824  
 Db 1558 EGSPFCLCNPLTLDLSGQ---RCVNTTSTB--DFPDHDIMDICMKKVTDVCSQ 1610  
 Cy 825 YPGNYPANTEC 835  
 Db 1611 LRGHHTYTEC 1621

## RESULT 10

US-09-230-652-2  
 / Sequence 2, Application US/09230652A  
 / Patent No. 6537775  
 / GENERAL INFORMATION:  
 / APPLICANT: Tournier-Lasserre, Elisabeth  
 / APPLICANT: Bousset, Marie-Germaine  
 / APPLICANT: Bach, Jean-Francois  
 / TITLE OF INVENTION: GENE INVOLVED IN CADASIL, METHOD OF DIAGNOSIS AND  
 / FILE OF INVENTION: THERAPEUTIC APPLICATION  
 / FILE REFERENCE: 03115.0048-00000  
 / CURRENT APPLICATION NUMBER: US/09/230.652A  
 / EARLIER FILING DATE: 1999-05-17  
 / EARLIER APPLICATION NUMBER: FR 96 09733  
 / EARLIER FILING DATE: 1996-06-01  
 / EARLIER APPLICATION NUMBER: FR 97 04680  
 / EARLIER FILING DATE: 1997-04-16  
 / EARLIER APPLICATION NUMBER: PCT/FR97/01433  
 / EARLIER FILING DATE: 1997-07-31  
 / NUMBER OF SEQ ID NOS: 163  
 / SOFTWARE: Patent In Ver. 2.1  
 / SEQ ID NO 2  
 / LENGTH: 2121  
 / TYPE: PRT  
 / ORGANISM: Homo sapiens  
 / FEATURE:  
 / OTHER INFORMATION: human ADNC No. 6537775ch 3  
 / US-09-230-652-2

Query Match 9.4%; Score 526.5; DB 4; Length 2121;  
 Best Local Similarity 24.2%; Pred. No. 3.8e-29;  
 Matches 232; Conservative 75; Mismatches 303; Indels 347; Gaps 58;

Cy 33 PFG-RGRAGGCEVDECAQGLDDCHADALCQNPRTYKSCSRGIVG----- 73  
 Db 147 PPGYQGRSC-RSDVDECRVG-EPCRHGGTLLNTPGSRQCPAGYTGPLCEPAVPCAR 203  
 Cy 80 -----EGRQCE-DIDECGNE--LNGGCVHGCNIPNRYAC 111  
 Db 204 SPGRNGGTCRQSSULTYDCAALPPEEGECNEVWDCGPHRCGLNGGTCVAVN---TNG 260  
 Cy 112 TCFGFMLAHJGHNCL-DVDECLENNAGCGH--TCVAVGSEYECCKEGF---FLSDN-- 163  
 Db 261 QCPPEW---TGCPCTEDVDECCQAPACHNGGTCFNTLGGHSCVAVNWTGSGSQND 316

Cy 164 ---ORTC:R-----SSEELGSCNKHGCSHICXE-----APRGVA 197  
 Db 317 DCAVAFCHGATCHDRVASFYCACPMKLTLLHLHDACVSNCHEDA:CTTPVNGRAI 376  
 Cy 138 GCEPFGELANQRCDCITCNHNGGCGH--SCDDTADGECSCGHPY--KMHTRGSC 252  
 Db 377 CTCPFGFGGACDQ-VDECSIGANPEHLGRCVNITGSPFLCCCRGYTGRCETDVNEC 435  
 Cy 253 J-----ERDYLVEITENNTSVVDGDRVAKRLXETCAN---NGCCETTCQDTSTGVH 305  
 Db 436 LSGCRNATCLDRICGFTGICVAGFTGYCEVDIDECQSSPVNGS---VCKDRVNGFS 492  
 Cy 306 CSCPFGTLQDGCTCK-CEDECCT---RNGGDHFENKIVASFQCGCKKFK-LJDEK 360  
 Db 493 CTCGSGFS---STCQLVDDECASTCRGSA---KVVDPPDGYECCAGAFEGCTLCD-- 543  
 Cy 361 SCQVDDECSLDRITCDHS-CINHPGTAFACNRYTLYGFTHC----- 401  
 Db 544 --RVVDDSPD-PCMHGRGVVDGLASFSCACAPGT--GTRCESQVDECRSPQRHGGK 597  
 Cy 402 -----GCTN-ECSINNGGCGC-----VCNVTGSEYECQCHGYKLNHMKK 441  
 Db 598 LDLYDKYLCGSGSTTGNGCEVNI:DCASNPCTPFGVGRDGINRDCVQCF-- 645  
 Cy 442 CVEVKGGLPITSVSPFVSLHCGKSGG--GGGCF-RCHSGIHLSSDVTITRTSVFKLNEG 498  
 Db 650 ---CGPLCNVEINECASPCGEGGCVDGSENGF-RC----- 681  
 Cy 499 KCSLKNALFPFEGLRP--ALPEKSSVKSSEFRVNLTCSSG--KQVPGARGSTPKEMF 554  
 Db 682 -----LCPGGLPPLCLPSPHCAHE-----PCHGICVDAFAG-- 715  
 Cy 555 ITVEBELETNKEVTASODLSC:VKRTENKLRKAI:RLRAVHREDFHLQLSGMNLDAVAK 614  
 Db 716 -----FRCV-----CEPGW 724  
 Cy 615 KPRTSERQAESCGVGGHAENOCVS--CAG--TYVDGAREC:LCPNSTFGNEEGM- 669  
 Db 725 SGPRCS-----SLIARDACESQPCRAAGTCSDDGAGFHC-TCPPAV-QGRQCELL 772  
 Cy 670 -TCBPCP-----RPGSGALKTPEANMMSCGGLQCPGEASAGTAPCCALGTG 719  
 Db 773 SPCTPKNPEHSGRESAPGC:PVSCCPQGC---GPRCCQVDECCAGPAPCG--PHGIC 926  
 Cy 720 QPEAGRTSCFPCCGGGLATKQGA-SFQDCETRVCCSPGHFNNTTHRCIRCFVGYQPEF 779  
 Db 927 TNLGSRSC-TCGG-----YTGSPCDDGIN---DCDNPCLNG-----GSCQGV 868  
 Cy 780 GKXNCVSCPGNTTIDPDG---STN:TOCKNRRCGGVJSDFTGYIES-----PNYFG 927  
 Db 869 GSFSCSCLPG-----FAGPRCARVDDELCLNPGCF--GTCTDVASFTCTCPPOYGG 918

## RESULT 11

US-09-185-412-19  
 / Sequence 19, Application US/08185432  
 / Patent No. 5750652  
 / GENERAL INFORMATION:  
 / APPLICANT: Ariavants-Tsakonas, Spyridon  
 / APPLICANT: Bousset, Isabelle  
 / APPLICANT: Diederich, Robert J.  
 / APPLICANT: Xu, Tian  
 / TITLE OF INVENTION: DELTA PROTEINS, NUCLEIC ACIDS, AND  
 / NUMBER OF SEQUENCES: 23  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSEE: PENNIE & SPOONDS  
 / STREET: 1155 Avenue of the Americas  
 / CITY: New York  
 / STATE: New York  
 / COUNTRY: U.S.A.  
 / ZIP: 10036-2711



```

Db 478 QGTFGRCEYIN:NECESHP---CQNEG-SCLDDEGTFFCVMPGFTG:GCGEIL:DS----- 529
Qy 280 RLMEYCAVNNOGCDRTCKJTS:GVHSCPVGFLLQDLGKTC-DIDECQT---RNSGCD 338
Db 530 -----CQSNPCLNGTCHDKIN:FKSCALGFT-----GACQINIDCGCPORNRNG-- 577
Qy 336 HFCXNIVGSPDCGCKKFKLLTDEKSCQ-DVDECSLDRTCHSHCIN:HGTFACACNRGYT 394
Db 578 -ICHDSIAGVSCCEPPGY-----TGTSCEIN:INDCSNPHRGKCIDVNSFK:CDPGYT 632
Qy 395 LY-----GFTHCQD-----TNEG:IN--NGCC 424
Db 633 GYICQKQINECESNPOCFDQ--HCDRVGSIYCCQCAQTSKNCENVNNECHSPCNNGA 690
Qy 415 QQCVNATVGSYECQCHPGYK-LHMNKKDCVEVKGLLPTSVSPRYSLHCGKSG-----GG 467
Db 692 --TCIDGINSYKCKQCVPGFTGQHEK-----NVDECSPPCANNGVCTIQVNG 736
Qy 468 DGCFLRCHSGI--HLSSDVTITRTSVTFKLINEKCS:KNAELF--PEGLR----- 513
Db 737 YKC--ECPRGFYDAHCLSDVDEGCASNPC--VNEGRCEGGINFICHCPGYT:KRCLELI 792
Qy 514 ---PALPEKSHSVSESFRVNLITCSSGQVPGAPGRPRTPKEMETIVSELETNCKE-VT 569
Db 793 DECSNPOQHOGT---CYDKLNAFSCQCMPTGYTGO-----KCEINIDCV: 935
Qy 570 ASCDL--SCIYKRETEKRLKRAIRTLKAVHREQFHLQSGNLDVAKP--PRTSERQAE 625
Db 836 NPGNGGTCIDK-----VNGYKC-VCKPFTFGRCESKMD 869
Qy 626 SCGVGQGHAEVQCV-----SCR---AGTYDGARERCIL---CPNG-TFQNEBQ 668
Db 870 PCARNRCKEAKCTPSSNFDLPSCCTCKLGYGRYCDDEDIDECSSSPCANASCLNVPGS 929
Qy 669 MTCEFCPR--PENSALKTPEAMNMBEGSGU-CQGEYSADGAPACQCALGTQPEAGR 725
Db 930 YRC-LCTKGYEGRDCAI-----NTDDCASFPQCGRTCCDGI-----GD 967
Qy 726 TSCFPGGGLATKHGATSPQDCETRVQ-----CSPG---HFYNTTHRCICPVQTY 775
Db 968 YSCU-CVDPFGDKH-----CETDINECLSGPQNGATSCQVNSYT---CICPELG-- 1033
Qy 776 QPEFGKNNVCSPGNTTDFDQSTYITQCKARRC--GGEJGP-FTGYIES--PYFGNYP 833
Db 1014 ---PFGINC-----QTMDEDCITESGCMGSCIDGINGVCSCLAGYSG--- 1054
Qy 831 ANTECTWTIN 840
Db 1055 --ANCCYKLN 1062

```

```

RESULT 13
US-08-185-432-18
Sequence 18, Application US/08185432
Patent No. 5750652
GENERAL INFORMATION:
APPLICANT: Artavanis-Isakoras, Spyridon
APPLICANT: Bussseau, Isabelle
APPLICANT: Diederich, Robert J.
APPLICANT: Xu, Tian
APPLICANT: Matsuno, Kenji
TITLE OF INVENTION: DELTAEX PROTEINS, NUCLEIC ACIDS, AND
TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSES: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:

```

MEDIUM TYPE: floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/185,432  
 FILING DATE: 21-JAN-1994  
 CLASSIFICATION: 536  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Wisstock, S. Leslie  
 REGISTRATION NUMBER: 18,672  
 REFERENCE/DOCKET NUMBER: 7326-006  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 869-8664/9741  
 TELEFAX: (212) 790-9090  
 TELEX: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 18:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2523 amino acids  
 TYPE: amino acid  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 US-08-185-432-18

Query Match 8.3%; Score 501.5; DB 1; Length 2523;  
 Best Local Similarity 23.0%; Pred. No. 2,6e-27;  
 Matches 233; Conservative 98; Mismatches 300; Indels 383; Gaps 65;

```

Qy 33 PGRGRARP--QEDVDECAQGLDDCHADALCONTPTSYKSCSKRPYQGEGRCE-DIDE 89
Db 400 PPS---YTPACNNDVDECSLGANPEHGRGRCNTILGSCQCNPPQYAP--RCELDVNE 454
Qy 90 CGNELNGGCVHD--CLNIPGNVRCCTCFDGMALHDGNC-LDVECTLENNGSCQHT--CV 144
Db 455 C---LSNPQNDCTCLDQIEFCQICMPGY---EGLYETNIDECASN--PCLHNGKCI 505
Qy 145 NIMGSYECCKEKEFFISDNQHNCIRSEB-----GLSCMKKHGCSHICKAEARG-- 194
Db 506 DKINERFCQPTGF--SGNLCQHDFTDCTSTPCXNGAKCLDGPSPSYTCCCTEGTGGH 562
Qy 195 -----SVACEGRPGE-----YAK---NORDC----- 213
Db 562 CEODINECIPDPCHVGTCDGATFTCLCPREYTGALCENDINECSKCLNQGCTDRE 621
Qy 214 ---ILCNHANGG--CQASCDTA-----DGPECSHPQYKXHTDGRSCLER 255
Db 622 NGYICCPKGTGTGNCEYK:DDCASRLCDNGKCIDXIDGVECTCEPGY---TGKLC-- 674
Qy 256 EDTVLEVTESNTTSVDDQKRYKRLMTCAVNNGGCR-----TCK 298
Db 675 --NININECDSPCRNGG-TCKDQ:NGFTVCYCPDGYDHMKCLSVNECNSNPCHGACH 730
Qy 299 DSTGVHSCPVGFTQLDGTCKD--DECOFR--NSGCHFFCN:VGSFPQCKKGF 353
Db 731 DGVNRYKCDCEAKMS-----GSNC-DJNNECESNPMNGT---CXYGAYICTGKAG 782
Qy 354 KLLTDEKSCQ-DVDECSLDRTCHS-----CINHF-- 392
Db 783 ---SGPNQTNINECSNPNLHSGCIDVAGYKNCMLPYTGA:CEAVLAPCAQSPCK 838
Qy 383 -----GTFACACNRGYTLVFTHOGCTNECS:VNGGQDQ--VGVNATVGSYECQ 429
Db 839 NGRCKESEDFFETFSCEPPGM--QQTCEIDNBEV--NRPKNGATCQNTNGSKNC 894
Qy 430 HPGYKJHNNKDC-VEVKG:LP:SVSPRYSLHCGKSGGDCGFLRCHSGIHLSSDVTTR 498
Db 895 KPGY-----TSRNEEMIDDOQPNP-----CHNG----- 918
Qy 489 TSVTFKLINEKCS:KNAELF--PEGLRPALPEKSHSVKESFRVNLITCSSGKY-- 540
Db 919 -----GSCS-DGIMNFFCNCPPGFRG--PKCELDINEC--ASNPKNGANCTDCY 963
Qy 541 -----PAPORPRTPKEMETIVSELETNCKEVTASGCDJSCYKREKRLKRAIRTL 592

```

```

Db 964 NSYTCTCCQGFSG-----IHESNTPTCTES---SCNGST----- 996
Qy 593 KRAVHREPHLOLQSMNLDVAKKPRTRSERQAESGCVGGHAENOCYS--CRAFTYYDGA 650
Db 997 -----CIDGINFTCCCPPTGTSYCO-----HDINECSKPLNG----- 1032
Qy 651 RERCILCPNGTFONEGQMTCEPCRPONSGALKTPKPAW-NMSEC--GGLOQPEYSADG 707
Db 1033 -----GTCCDSYGTYYKC--TCPOGYTGALNCQULVWRCDSSPCANGKC-----WCTIN 1278
Qy 708 FAPCOLCALG--TFQPEAGRTSCFPGGGLATKHOGATSPQDCETRVQCSGPHFYNTTH 765
Db 1079 FYRCE-CKSGMTGVYCDVPSVSC-----EVAKKQGVVDIYHLCRNSGMCVD---TGNTH 1128
Qy 766 RCIRCPVGTIPEFGKNCVSCP-----GNTTDFDG-----STNITQC 804
Db 1129 FC-RCQAG-YTGSYCEQVDECSFNPQNGATCTDYLGSYCECVAGYGVNCSEINEC 1186
Qy 805 KMRRC--GGELGDFGTGYIESPNYPNPANT-----ECT-----WTINP 841
Db 1187 LSHPCQNGSTCIDJLN-----TYKSCPRGTGCHCEINVDCTPFYDSFTLEP 1235

```

```

RESULT 14
US-08-899-232-3
Sequence 3, Application US/08899232
Patent No. 6436650
GENERAL INFORMATION:
APPLICANT: Attavani's-Tsakonas, Spyridon
APPLICANT: Qi, Huijin
TITLE OF INVENTION: ACTIVATED FORMS OF NOTCH AND METHODS BASED THEREON
FILE REFERENCE: 7326-046
CURRENT APPLICATION NUMBER: US/08/899,232
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patent Ver. 2.0
LENGTH: 2523
SEQ ID NO 3
TYPE: PR
ORGANISM: Xeropus sp.
US-08-899-232-3

```

```

Query Match 8.9%; Score 501.5; DB 4; Length 2523;
Best Local Similarity 23.0%; Pred. No. 2,6e-27;
Matches 223; Conservative 98; Mismatches 300; Indels 383; Gaps 65;

```

```

Qy 33 PPGRGRAAP--QEDVDECAQGLDCHADALCQNTPTSYKSCKRGYQJEGRQCE-DIDE 99
Db 400 PPG---YTGPAQNNVDDECSLQANPEHGRCNTLGSFQCNCPQYAGP--RCEIDVNE 454
Qy 90 CGNELNGGCVHD--CLINPGNYRCTGFGFMAHGHNC-LVDECLENNGCCQHT--CY 144
Db 455 C--LSNPQNDSTCTDQIGFPOQICMPY---BGLYETIYIDCASH--PCLHNGKCI 505
Qy 145 NWMGYECCCKEGLFSDNQHTCIHRSE-----GLSCMKKHGSHLCKEAPR3-- 194
Db 506 DKINEFRCDPTGF---SGNLQGHDFDCTSPCKNGAKCLDGNSTYCCQCTEFTFRH 561
Qy 195 -----SVACGRPGFE-----LAK--NQKDC----- 213
Db 562 CQDINECIIPDCHYGTCKDG:ATFTCLCRPGYGTGLCNDINECLSKCLNGGQCTBRE 621
Qy 214 ---ILTCNNGNG--CQHSQCDPTA-----DQPECSCHPOYMHNDGSCLER 255
Db 622 NGYICTCPKGTGTCNGSTKIDDCASNLCKNGKIDKIDYECTCTBERGY---GKLC--- 674
Qy 256 EDTVLEVTESNTSYVDGDKRYKRLMETCAVNNGGCDR-----TCX 298
Db 675 ---NININCDSPKRNCG--TCXQDINGFTVCVCPDGYHDMCLSEVNECNSPCHGACH 730
Qy 299 DSTGVHSGCPVGFITQJLQKTKCD--DEGCTR---NGSCHPCKN:VGSDDCCCKXGF 353

```

```

Db 731 DGVMGYKCDCEAGWS-----GSNC-DINNNECESNPQNMGT---CKMTGAVICTCKAGF 782
Qy 354 KULTDEXSCQ-DVDECSLDRCTDHS-----CINHP-- 382
Db 783 ---SGPNCQTNINECSNPCLNHGCTCIDVAGYCNCKMLPTGATICEAVLAPCAGSCK 838
Qy 383 -----GTFCACNRYTLYGTHGCDTNECSINMGCCQ--VCVNTVSGYEQC 429
Db 839 NGGRCKESSEDFEFSECEPPGW--QQQTEIIMNECV--NRPCRNATCQNTNGSYKCNK 894
Qy 430 HPGYKUHANKKC-VEWKLPTSVSPRVSLHCKSGGGGCGFLACHSHI:HSVTTIR 488
Db 895 KPGY-----TGRNCMDIDDCQPNP-----CHNG----- 918
Qy 489 TSVTFKLNBSKSLKALFP-----PEGALPAPKXSVKSEFRVNLTCSSKQV----- 540
Db 919 -----GSCS-DCIYMFPCNCPAGFRG--PKCEDINEC--ASNPCYKANGCTDVC 963
Qy 541 -----PGAPGRPSPKXEPITVEFELEITNOKEVTASCD:SCIVKATEKRLKRAITL 592
Db 964 NSYTCTCCQGFSG-----IHESNTPTCTES---SCFNGT----- 996
Qy 593 KRAVHREPHRLQSMNLDVAKKPRTRSERQAESGCVGGHAENOCYS--CRAFTYYDGA 650
Db 997 -----CIDGINFTCCCPPTGTSYCO-----HDINECSKPLNG----- 1032
Qy 651 RERCILCPNGTFONEGQMTCEPCRPONSGALKTPKPAW-NMSEC--GGLOQPEYSADG 707
Db 1033 -----GTCCDSYGTYYKC--TCPOGYTGALNCQULVWRCDSSPCANGKC-----WCTIN 1278
Qy 708 FAPCOLCALG--TFQPEAGRTSCFPGGGLATKHOGATSPQDCETRVQCSGPHFYNTTH 765
Db 1079 FYRCE-CKSGMTGVYCDVPSVSC-----EVAKKQGVVDIYHLCRNSGMCVD---TGNTH 1128
Qy 766 RCIRCPVGTIPEFGKNCVSCP-----GNTTDFDG-----STNITQC 804
Db 1129 FC-RCQAG-YTGSYCEQVDECSFNPQNGATCTDYLGSYCECVAGYGVNCSEINEC 1186
Qy 805 KMRRC--GGELGDFGTGYIESPNYPNPANT-----ECT-----WTINP 841
Db 1187 LSHPCQNGSTCIDJLN-----TYKSCPRGTGCHCEINVDCTPFYDSFTLEP 1235

```

```

RESULT 15
US-08-893-590A-20
Sequence 20, Application US/08083590A
Patent No. 5786158
GENERAL INFORMATION:
APPLICANT: Attavani's-Tsakonas, S. et al.
TITLE OF INVENTION: Therapeutic And Diagnostic Methods
TITLE OF INVENTION: And Compositions Based On No. 5786159ach Proteins And
NUMBER OF SEQUENCES: 2;
CORRESPONDENCE ADDRESS:
ADDRESSES: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/083,590A
FILING DATE: 25-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mierock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7326-015

```

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212 780-9090  
 TELEFAX: 212 8698864/9741  
 TELE: 66141 PENNIE  
 INFORMATION FOR SEQ ID NO: 20:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2556 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: unknown  
 MOLECULE TYPE: peptide  
 US-08-093-590A-20

Query Match 8.8%; Score 491.5; DE 1; Length 2556;  
 Best Local Similarity 23.7%; Pred. No. 1.4e-26;  
 Matches 238; Conservative 90; Mismatches 339; Indels 337; Gaps 5;

```

QY 41 GP--QEDVDECAAGLDGCHADALCONPTSYKCSCK----- 74
Db 172 GPTCRQDVNECGQKRLCRHGGTCHNEVGSYRCVCRATHTGFCNCRPYVPSPSPCCNGS 231
QY 75 -----PSYQGEGRQCEIDIEC--GNEI--NGGCVHDCLN--PQNYRCTCFDGF 117
Db 232 TCRPTGVTHECACLPFTGQNC--ENIDCPGNCKKNGACVGVN---TYNCPCEPW 287
QY 118 MLAHDGHNCL-DYNEBCEJENGGCQH--TCVVMGSEYECCKEGRFF---LSDN-----QHT 165
Db 288 ----TGQYCTEDVDECOQLMPNACONGTCHNTHGYNCCVNVGWTGEDCSENIDDCASAA 343
QY 167 CIH-----RSEGLSCMKNDHGCSECKEAP-----RGSVACECRPG 203
Db 344 CFHGTCHDRVASFCEPHGTGLCHLNDACISNFCNESSNCDTNPVNGKALCTCPSG 403
QY 204 FELAKNORCIYTCNNGNGGCQHS--CDDIADGPECSCHPY---KMTDGRSCL---E 254
Db 404 YTGPAQSGD-VDECSLGANPCEHAKCINTLGSFECCQLGTYGPRCEIDVNECVSNPCQ 462
QY 255 REDTYLEVTESENTSVVDGRVVRKRLMETCA----VNNGGCDRTCKDTSTGVHSCPV 310
Db 463 NDATCLDQIGFQCMCKPFGYEVHCEVNTDECASSPCLHNG----RCLDKINEFQCECFT 518
QY 311 GFTLQDQKCTCK-DIDEQCT---RNGGCDHFCRN--VGSFDCGKKGFALLTDEKSCDDVD 366
Db 519 GFT---GHLQOYDVDECASTPCKNGA---KCLDGRNTYTCVCTBGY---TGTHCEVDID 568
QY 367 ECSLDRTCDHSCINRGTFACACNRYTYVGFTHC-GDTNECSIN---NGGCOQ----- 416
Db 569 ECDPDPCHYGSKDKDVATFTCLCRGYTGH---HCEFTNINECSSQPCSLRGTCQDPNAY 625
QY 417 -----VCVNTVGSYECQCHPGYKLMNKKDCVEVKGL 448
Db 626 LCFCLKTGTGPNCEINLDDCASSPCDSSGTCJDKITDGYECACEPGTJSMONSNIDECAG- 684
QY 449 LPTSVSPVSLHCGKSG---GGDGCFLRCHSGIHLSSCVTITRTSVTFKLINEKCSILXN 524
Db 685 ----NF-----CHMGTCDEGIDINFTCRCEGYH---DPTCLS-----EYNE--CN--- 721
QY 505 AELPPEGLRPALEPHKSSVKESFRVNLTCSSGKOVFGAPGRPSTPKEMFTTVEFELETN 564
Db 722 -----SNFCVHAGACRDLNSYKCCD-----PCMSG----- 747
QY 565 QKEVTASCDLS-----CIYKRTFKRLKAKITRLKAVHRECFHJOLSGMNLVAKKP 616
Db 748 -----TNCIDINNCESENPQVNGGTCKDMTSGI---VCTGREGF---SGPN----- 787
QY 617 FRTSERAESGCGVQGHAEHQVSCRAGTYYD--GAREPILCENGTFQNEGOMTCE- 672
Db 788 ---CQTNINEC-----ASNPCIN--KGTCIDVAGYKNCCLPYTGA-----TCBY 826
QY 673 ---PC-PRP-GNSGALKTPBAMNMSGC-----GLCQPGSEISADGFADCCLOALGTF 715
Db 829 VLAFCAPSPFCRNGSECESEYSEFSVCPTAGAKQTCCEV-DINECVLSPCRHGA----- 883

```

```

QY 726 QPEAKRSCFPCCGG-ATKHQATSPDCCEPRV-QCSPGHFY-----NTTHRCIR 769
Db 884 -----SCQNTHGYSRCHQAGYSGRNCBETIDDCRPNPCHNGSGSCDDGINTATCDLP 936
QY 770 CPVGYQPEFCRKNQVSCP---GNTTID-----FQZ---STNITQCKNRRC-- 809
Db 937 GFRGTFCBE-D-NECASDPGRNAGNCTDCVDSYTCCTGAGFSGHCENTPCTESSCFN 995
QY 310 GSE---LGDFTGYIESPNYPGNYPANTTECTWTINPPPRR-LI 849
Db 996 GGTGVAGINSFT-CLCPPGFTSSY-----CQHVAVECDSCRPLD 1033

```

Search completed: October 22, 2003, 11:31:49  
 Cdb time : 39 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OX protein - protein search, using sw model

Run on: October 22, 2003, 11:24:13 ; Search time 121 Seconds

(without alignments)  
2130.533 Million cell updates/sec

Title: US-09-747-371-2

Perfect score: 5635

Sequence: 1 MGAGNRNRGAAMAVLLLL.....RRSF-RLRSKSRFLRPYK 999

Scoring table: BLOSUM62

Gapcp 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp Vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteria:\*
- 17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	5635	100.0	999	4 Q9NQ36	Q9NQ36 homo sapien
2	5056	90.2	997	11 Q9JUS0	Q9JUS0 mus musculu
3	3662.5	65.3	988	4 Q81MY4	Q81MY4 homo sapien
4	3640.5	65.0	992	11 Q8C9Q4	Q8C9Q4 mus musculu
5	3397	60.6	993	4 Q8IX30	Q8IX30 homo sapien
6	3262.5	58.2	961	11 Q9EQC6	Q9EQC6 mus musculu
7	2916	52.0	880	4 Q8NAU9	Q8NAU9 mus musculu
8	1738.5	31.0	581	4 Q8NAV8	Q8NAV8 mus musculu
9	875	15.6	1664	5 Q9TYQ2	Q9TYQ2 caenorhabdi
10	841.5	15.0	1574	11 Q88281	Q88281 ratu
11	775.5	13.8	1463	5 Q9V589	Q9V589 dirosophila
12	737.5	13.2	1246	4 Q75095	Q75095 mus sapien
13	720	12.8	2906	11 Q9WU89	Q9WU89 ratu
14	713.5	12.7	2893	4 Q96TF8	Q96TF8 mus sapien
15	695	12.4	810	11 Q8C435	Q8C435 mus musculu
16	695	12.4	956	11 Q99K64	Q99K64 mus musculu

17	695	12.4	956	11 Q8R542	Q8R542 mus musculu
18	695	12.4	1289	5 Q8SSS3	Q8SSS3 dictyostei
19	694	12.4	2189	5 Q9B105	Q9B105 elmeria ten
20	693	12.4	3857	11 Q88840	Q88840 mus musculu
21	686.5	12.2	937	4 Q96FT5	Q96FT5 homo sapien
22	686.5	12.2	1016	4 Q8NDE6	Q8NDE6 homo sapien
23	675	12.0	2872	11 Q9WU88	Q9WU88 ratu
24	654.5	11.7	61	4 Q9Y559	Q9Y559 mus sapien
25	640	11.4	708	13 P87363	P87363 gallu
26	610.5	10.9	1062	11 Q60789	Q60789 mus musculu
27	610	10.9	1398	13 Q9AXM6	Q9AXM6 xenopus lae
28	607	10.8	1399	13 Q8JF24	Q8JF24 xenopus lae
29	601	10.7	1389	11 Q8CG18	Q8CG18 mus musculu
30	601	10.7	1713	11 Q8CG19	Q8CG19 mus musculu
31	598	10.7	1713	11 Q88349	Q88349 mus musculu
32	597.5	10.7	1963	6 Q28319	Q28319 bos tauru
33	587	10.5	1382	4 Q9HFK2	Q9HFK2 homo sapien
34	587	10.5	1972	5 Q8WPL1	Q8WPL1 oikopleura
35	583.5	10.4	1764	11 Q35806	Q35806 ratu
36	578	10.3	293	6 Q8WJK0	Q8WJK0 cercopithe
37	575.5	10.3	733	5 Q25678	Q25678 podocoryne
38	575	10.3	1551	5 Q9NGV4	Q9NGV4 drosophi
39	575	10.3	3396	5 Q9VM55	Q9VM55 dirosophila
40	574	10.2	1174	11 Q95K58	Q95K58 mus musculu
41	571	10.2	1253	11 Q61810	Q61810 mus musculu
42	570.5	10.2	1821	4 Q14767	Q14767 homo sapien
43	567	10.1	658	4 Q8NB66	Q8NB66 homo sapien
44	567	10.1	685	11 Q922K8	Q922K8 mus musculu
45	566.5	10.1	746	4 Q96HB9	Q96HB9 homo sapien

#### ALIGNMENTS

RESULT :

ID Q9NQ36 PRELIMINARY: PRT; 999 AA.

AC Q9NQ36; 31-OCT-2000 (TREMBLrel: 15, Created: 31-OCT-2000 (TREMBLrel: 15, Last sequence update: 31-MAR-2003 (TREMBLrel: 23, Last annotation update: CEGPI protein.

GN CEGPI.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Bahr A., Hankeln T., Winterpacht A., Zabel B., Schmidt E.R.; Comparative sequencing of Human chromosome 1p15 and mouse chromosome 7.

RT Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: CONTAINS 1 CUB DOMAIN.

DR HSSP; P07204; IADX.

DR InterPro; IPR000152; Asx hydroxyl.

DR InterPro; IPR000659; CUB domain.

DR InterPro; IPR001881; EGF\_CA.

DR InterPro; IPR006209; EGF-like.

DR Pfam; PF00431; CUB; 1.

DR Pfam; PF00009; EGF; 9.

DR SMART; SMO042; CUB; 1.

DR SMART; SMO0179; EGF\_CA; 6.

DR PROSITE; PS0010; ASX\_HYDROXYL; 6.

DR PROSITE; PS01180; CUB; 1.

DR PROSITE; PS01186; EGF\_2; 8.

DR PROSITE; PS01187; EGF\_CA; 6.

KA EGF-like domain.

SC SEQUENCE 999 AA; 109956 MW; 61314844A053095 CRC64;

Query Match 100.0%; Score 5635; DB 4; Length 999;

Best Local Similarity 100.0%; Pred. No. 0;

Matches	999	Conservative	0	Mismatches	0	Indels	0	Gaps	0
QY	1	MGVAGNRPGAAAVLLILLPLP.LLLAGAVPPGRGAAAGPQEDVDECAQGLDDCHADA	60						
DB	1	MGVARNRPGAAAVLLILLPLP.LLLAGAVPPGRGAAAGPQEDVDECAQGLDDCHADA	60						
QY	61	LCQNTPTSYKSCSKPGYQGEGRQCEIDIDECGNELNGGCVHDCLNIPGNYRCTCFDGFMLA	120						
DB	61	LCQNTPTSYKSCSKPGYQGEGRQCEIDIDECGNELNGGCVHDCLNIPGNYRCTCFDGFMLA	120						
QY	121	HDGHNCLDVEDECLENNGGCHTCVNMVMSYECCKEGFF.S.DNCHTCHRESEGLSCMNK	190						
DB	121	HDGHNCLDVEDECLENNGGCHTCVNMVMSYECCKEGFF.S.DNCHTCHRESEGLSCMNK	190						
QY	181	DHGCSHCKEAPRGVACCEPRGFELANQDCLITCNHNGNGGCHSCDDTADGPECSCCH	240						
DB	181	DHGCSHCKEAPRGVACCEPRGFELANQDCLITCNHNGNGGCHSCDDTADGPECSCCH	240						
QY	241	POYKMTDGRSCLEREDTVLEVTESNTTSVVDGDKRVKRLJMETCAVNNGGCDRTCKDT	300						
DB	241	POYKMTDGRSCLEREDTVLEVTESNTTSVVDGDKRVKRLJMETCAVNNGGCDRTCKDT	300						
QY	301	STGVHSCPVGFTLQDGTCKDIDECOTRNGGCHPFCKNTVGSFDDGCKKGFLLTDEK	360						
DB	301	STGVHSCPVGFTLQDGTCKDIDECOTRNGGCHPFCKNTVGSFDDGCKKGFLLTDEK	360						
QY	361	SCQDVDECSLDRTCDSHCINHPGFACACNRGYTLVGFTHCGDITNECSINNCGCQVCYN	420						
DB	361	SCQDVDECSLDRTCDSHCINHPGFACACNRGYTLVGFTHCGDITNECSINNCGCQVCYN	420						
QY	421	TVGSYECQCHPGYKLMNKKDCVEVYKGLPTSVSPVSLHCKSGSGGCGFLRCHSGIHL	480						
DB	421	TVGSYECQCHPGYKLMNKKDCVEVYKGLPTSVSPVSLHCKSGSGGCGFLRCHSGIHL	480						
QY	481	SSDVTTIRTSVTFKLNKSGCSLKNALFPEGRLPALPEKHSVKESPRYVNLTCSSGKQV	540						
DB	481	SSDVTTIRTSVTFKLNKSGCSLKNALFPEGRLPALPEKHSVKESPRYVNLTCSSGKQV	540						
QY	541	PGAPGRSTPKEMFTVFELETNQEVYASCDLSICVKTREKRLKKAIRTLRAVAREQ	600						
DB	541	PGAPGRSTPKEMFTVFELETNQEVYASCDLSICVKTREKRLKKAIRTLRAVAREQ	600						
QY	601	FHLQSGNLDVAKKPPRTSERQASCSVGGCHAHNCQVSCRAGTYVYAGARECCLCNG	660						
DB	601	FHLQSGNLDVAKKPPRTSERQASCSVGGCHAHNCQVSCRAGTYVYAGARECCLCNG	660						
QY	661	TFQNEGQWTCCEPRPGNSGALKTPPEAMNMECCGGLCOPGBYSADGAFQCLALSTFQ	720						
DB	661	TFQNEGQWTCCEPRPGNSGALKTPPEAMNMECCGGLCOPGBYSADGAFQCLALSTFQ	720						
QY	721	PEAGRTSCFPGGGLATYHQAATSFQDCETRYQSGFHYNTTHRCIRACVGTYYQPEFG	780						
DB	721	PEAGRTSCFPGGGLATYHQAATSFQDCETRYQSGFHYNTTHRCIRACVGTYYQPEFG	780						
QY	781	KXNVSCPGNTTDDGSTNTTQCKNRGCGELGFTGYIESPRYKPYNTECTETVYN	840						
DB	781	KXNVSCPGNTTDDGSTNTTQCKNRGCGELGFTGYIESPRYKPYNTECTETVYN	840						
QY	841	PPKRRILIVPEILFIEDDCGDTLNMKTSSSVTTEFCQYERPAFTSRSKLM	900						
DB	841	PPKRRILIVPEILFIEDDCGDTLNMKTSSSVTTEFCQYERPAFTSRSKLM	900						
QY	901	LOKSNENGSARGFVYVYDEDEYQELIEDIVRDGRLYASENHQELKDKKLKALEDV	960						
DB	901	LOKSNENGSARGFVYVYDEDEYQELIEDIVRDGRLYASENHQELKDKKLKALEDV	960						
QY	961	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	961	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
QY	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						
DB	999	LAHPGNYKTYACSEREMPPSEFILLSKVSRFLRPYK	999						

```

Db 479 SSDVTVRTSVTFKINECKSIQAKLSPEGLRALPERHSSVESFOYANTLTSFGKOV 538
Qy 541 PGAPRSTPKEMFTFEFELTNOKEVTASCDLSICIVKTEKRJRAIRLRAVAREQ 600
Db 559 PGLALRLNAPKEMFTFEFERETKEKVTASCNISCVKRTKRLALRLTKAAAREQ 598
Qy 601 FHLDSGNLDVAKKPRPTSEROASCGVGGGHAENOCVSCRACTYYDGAERECILCPNG 660
Db 599 FHLDSGMDLMAKTPRSVSGOHEETCGVGGHESQCVSCRACTYYDGSERCILCPNG 658
Qy 661 TFONEBQMTCEPCRPCNSGALKTPPAMNMSDEGGLCQPEEYSADGFAPQOLCALSTFC 720
Db 659 TFONEBQVTCPCRPCPNLGLSKISEAMNVSDCGGLCQPEEYSANGAPQOLCALSTFC 718
Qy 721 PEAGRTSCFPGGGGATGAGATSPQOCETRVQSPGFHYTTTHRCRCRCVGTGYGPEFG 730
Db 719 PDVGRITCLSCGGGLPTHLGATSPQOCETRVQSPGFHYTTTHRCRCRCVGTGYGPEFG 778
Qy 781 KNCVSCPGNTTTFDGSNTITQCKNRRCGELDFTGYESPNYPGNYPANTECTWTIN 840
Db 779 KNCVSCPGNTTTFDGSNTITQCKNRRCGELDFTGYESPNYPGNYPANTECTWTIN 838
Qy 841 PPKRRLIIVPEITFLIEDCGDYLVRKTSNSNSVTYTCQTYERPIAFTRSKLW 900
Db 839 PPKRRLIIVPEITFLIEDCGDYLVRKTSNSNSVTYTCQTYERPIAFTRSKLW 898
Qy 901 IOPKNEGNSARGFVPTVYDEDEYOELIEDIVRQRLYASENHQELIKKKLKALFDV 960
Db 899 IOPKNEGNSARGFVPTVYDEDEYOELIEDIVRQRLYASENHQELIKKKLKALFDV 958
Qy 961 LAHPQNYFYTAGESREMPFRSFRILRLSKVSRFLRPYK 999
Db 959 LAHPQNYFYTAGESREMPFRSFRILRLSKVSRFLRPYK 997

RESULT 3
ID 081WY4 PRELIMINARY; PRT; 988 AA.
AC 081WY4;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Signal peptide-CUB-EGF-like domain containing protein 1.
GN SCUBEL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_taxid=9606;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=2236413; PubMed=12270931;
RA Yang R.B., Ng C.K.D., Maeserum S.M., Colman S.D., Shenoy S.,
RA Mehrahan F., Komuves L.G., Tomlinson J.E., Topper J.N.;
RT "Identification of a Novel Family of Cell-surface Proteins Expressed
RT in Human Vascular Endothelium."
RL J. Bio. Chem. 277:46364-46373(2002).
DR EMBL; AF525689; AAN77133.1;
SQ SEQUENCE 988 AA; 107901 MW; 1BF57BEF786C9C1 CRC64;

Query Match 65.3%; Score 3662.5; DB 4; Length 988;
Best Local Similarity 62.2%; Pred. No. 2,9e-317;
Matches 652; Conservative 133; Mismatches 164; Indels 119; Gaps 14;

Qy 10 GAA---MAVLLTLLPLLLAGAVPPGRRAAG-----PQEDVDECAOGDPCCHADAL 61
Db 2 GAAVVRHLVLLAL-----GTRGLLAGSGGLPSGVVDEDSBGTDCHDAI 49
Qy 62 CNPTSYKSCCKPYOGEGROCEDIDECGNE-LNGSVHCLNIPGVYRGTCTDGFYLA 120
Db 50 CQNAKSYKCLCKPGYKGGKQCGDIDECNENYVGGCVHECINIPGVYRGTCTDGFYLA 109
Qy 12 HDGHNCLDVDEGLENNGGCQHTCVAVVMSYECCGKEGFFLSDNQHTCHSEBGLSCMK 160

```

```

Db 110 HDGHNCLDVDEEQDNNGGCQHTCVAVVMSYECCGKEGFFLSDNQHTCHSEBGLSCMK 169
Qy 181 DHGSGHYCKEARGSVACECRPGFELAKNORDCITCNHNGNGCQHSDDPTADGEGSC 240
Db 170 DHGSAHYCKETFKGVACDCRPGFLAQNQKCTITCNVNGGCGHSEDDTGTGCGH 229
Qy 241 POKKTDGRSCLEREDTVLEVTESNTTSVVDGDKRVKRLMETCAVNNGCDBRTCKDT 300
Db 230 OKYAPHSDBRTCL-----ETCAVNNGCDBRTCKDT 259
Qy 301 STGVHSCGVGFTLLDGKTCQDIBQCTRNCGCHFCKNIVGSPDCCGKSKFKLITDEK 360
Db 260 ATGVRCSCVGFLLQPDGKTCQDINECLVNNGCDFGANTGVSEFCGCRKMYKLTIDER 319
Qy 361 SCQDVDEGLDPTCHSCINHGTFACACNRYTLXGFHPCDTEBGSINNGCQGVN 420
Db 320 TQDDIDEGSEFKTCHICINSPGSFOCLHROYILYGTTHCGSDVECSKNSGDCQGVN 379
Qy 421 TVSYECQCHPGYKLMNKKDCVEV-KGLLPTSVSPRYSJHCKSGGSDGCFLRC----- 474
Db 380 TKSRYECVCPGRRLHMNRKDCVEFGKCLSRAKTSPPAQLSCKAGVESCFLSCPAHTL 439
Qy 475 -----HSGTHLS---SDVTTRTSTVFLNBECKSL 502
Db 440 FVPDSNSYVLSCVPPGQKALQKRNQTSGLGSCSDAPRTPIKQKARFKIRAKCHL 499
Qy 503 KNAELFPESCLRALPEKESVYKESR-----YNNLCSSGKQVGAAGRPTPK 552
Db 500 R-----FHSQRAKETRQPLCHCHVTFTVLCDSKSK--RRGRKSPSKE 544
Qy 553 V-FIVEPELEFETNOKEVTAACDLSCVKTERTRLKRAITLPKAVHSEQFHLQSGMND 611
Db 545 VSHITHEFIEFTMEASDQCEADCCRKRAEQLGALITLAKSLIORQFYGVGSTERE 604
Qy 612 VAKKPRTERQAESQGVGCGHAENOCVSCRGCTTYDGAERECILCPNKTFQNEBQMT 671
Db 605 VAKRPAKALEGOS-AGAGQVQLQDSKCVACGPTHTFGELGQCVCPQTYQDMEGQLSC 663
Qy 672 EPCPRRNGSGLKTEPAMKXSECGGLCQPEEYSADGFAPQOLCALSTFCPEPSCFPC 731
Db 664 TPCP---SSDGLCPKARVNSDEGGLCQPEEYSADGFAPQOLCALSTFCPEPSCFPC 723
Qy 732 GGSJATKAGATSPQOCETRVQSPGFHYTTTHRCIRCPVGTGYGPEFKNKNCVSCPGNT 791
Db 722 GGSJATKAGATSPQOCETRVQSPGFHYTTTHRCIRCPVGTGYGPEFKNKNCVSCPGNT 780
Qy 792 TTFDGSNTITQCKNRRCGELDFTGYESPNYPGNYPANTECTWTINPPKRLIIV 851
Db 781 STFDGSIWVTHCKQHQHCGELDFTGYESPNYPGNYPANTECTWTINPPKRLIIV 840
Qy 852 PEITFLIEDCGDYLVRKTSNSNSVTYTCQTYERPIAFTRSKLW IOPKNEGNSA 911
Db 841 PEITFLIEDCGDYLVRKTSNSNSVTYTCQTYERPIAFTRSKLW IOPKNEGNSA 900
Qy 912 RGFQVPTVYDEDEYOELIEDIVRQRLYASENHQELIKKKLKALFDVLAHPQNYFYK 971
Db 901 RGFQVPTVYDEDEYOELIEDIVRQRLYASENHQELIKKKLKALFDVLAHPQNYFYK 960
Qy 972 AOSREMPFRSFRILRLSKVSRFLRPYK 999
Db 961 AOSREMPFRSFRILRLSKVSRFLRPYK 988

RESULT 4
ID 08C9Q4 PRELIMINARY; PRT; 992 AA.
AC 08C9Q4;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Signal peptide.
OS Mus musculus (Mouse).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium.  
 RA The RIKEN Genome Expedition Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RL Nature 420:563-573(2002).  
 SR EXEL; AK041580; BAC09961.1;  
 SQ SEQUENCE 992 AA; 108570 XM; 42219D667979.TA1 CR664;

Query Match 65.0%; Score 3640.5; DB 1.; Length 992;  
 Best Local Similarity 63.3%; Pred. No. 2,763,35;  
 Matches 658; Conservative 103; Mismatches 180; Indels 99; Gaps 15;

QY 13 GAA---MAVJLLLLPPLLALAGAVPGRGAAAPQEDVDECAQGLDCHADALCCVTP 66  
 DB 2 GAAAVRMHLYLLLLALGARGLVGSGLPG-----AVDDECEGTDCHIDALCQNTP 54  
 QY 67 TSYKSCPKPGYQGEGRQCEIDIECGNE-LNGCVHDCANI PGNYRCTCFDGFMLAHGHN 125  
 DB 55 KSYKCLCPGKGGKQCEIDIECGNE-LNGCVHDCANI PGNYRCTCFDGFMLAHGHN 114  
 QY 126 CLDVDECLENNNGGQCHTCVVMVGSYECCEKEGFLSDNQHICHRSEGLSCMKKHCQS 185  
 DB 115 CLDVDECLENNNGGQCHTCVVMVGSYECCEKEGFLSDNQHICHRSEGLSCMKKHCQS 174  
 QY 186 HICKEAPRGSVACRPGFELAKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQK 245  
 DB 175 HICKEAPRGSVACRPGFELAKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQK 234  
 QY 246 HTDRSCLEDEDTLVETESNTSVVDKDKVKRRLXETCAVANNNGCDBTCKDTSRGVH 305  
 DB 235 HADGRCTI-----ETCAVANNNGCDBTCKDTSRGVH 264  
 QY 306 CSCPVGFTLQDGTCKDIDECQTRNGGCHFCNKIVGSPDQCKGKPKLLTDEKSCQV 355  
 DB 265 CSCPVGFTLQDGTCKDIDECQTRNGGCHFCNKIVGSPDQCKGKPKLLTDEKSCQV 324  
 QY 366 DECSLDRTCDHSCINHPCTFACNKGITLYGFTHCQDQKQKQKQKQKQKQKQKQKQK 425  
 DB 325 DECSLDRTCDHSCINHPCTFACNKGITLYGFTHCQDQKQKQKQKQKQKQKQKQKQK 384  
 QY 426 ECQCHPGYKLMNKKDCVEVKGGLPTS-VSPRVSJHCGKSGGQDGCFLRC----- 474  
 DB 385 ECQCHPGYKLMNKKDCVEVKGGLPTS-VSPRVSJHCGKSGGQDGCFLRC----- 444  
 QY 475 HSGIHLSSDV-----TTITSVTFKLNKSCS--KKA 525  
 DB 445 ESSYILSCGVPGGLQKTLPRKNNGSSSTGPGSCDAPCTTPIKQKARFKIPAKCHILQRSQ 504  
 QY 506 ELFPREGI-R-PAPEKHSVKSFRVYVLTGSSGQVPGAPBGRSTPREX-FITTEPLET 563  
 DB 505 ERAKDTCRAHPLDNCVHT-----FVTLKQDSKKT-RRRGRKSPSEVSHITNAEFVEM 556  
 QY 564 NQKEVTASCDLSCTVKTETKRLRAIRTLRAVHREGFHQLSGNLDVAKKPERTSERQ 623  
 DB 557 KYDEASGTCEADCMKRAEQLQAIKILRKISGRNQGVVGLTEYEVAGRPKALAGT 616  
 QY 624 AESCGVQGHAEQVSCRACTYIDGAEKRCILPNTGTFQNEBQVQTCPEPRGNSGA 683  
 DB 617 G-TGGIGQIQDQKGVPCAPGTYSFGDQGMPCVSGTYQVDEBQISTPTP--SSSECL 672  
 QY 684 KTPKAMNSCEGGLQCPGEYASGAPFQCALGTFQPEAGRTSCFPQGGSLATKRGAT 743  
 DB 673 GLAGARNVSECQGGCSFGYFADGFKFQACPVGTIQPEPRGTGCFQGGSLATKRGAT 732  
 QY 744 SPQDCETRVQSGPGHYVNTTTRICIRCPVGTYQPEFGKNNVSCPGNTTTFDQSTNITQ 803

DB 733 SPQDCETRVQSGPGHYVNTTTRICIRCPVGTYQPEFGKNNVSCPGNTTTFDQSTNITQ 792  
 QY 804 CNRRQCELOCPFYGIESPPYPPANPANTCTWTINPPKRLIIVPEFLPIEDDCG 863  
 DB 793 CNRRQCELOCPFYGIESPPYPPANPANTCTWTINPPKRLIIVPEFLPIEDDCG 852  
 QY 864 DYLVRKRTSSNSVYTYCTCTYERPIAFTRSKKLWQFNSNNGNSARQGVYVYTD- 922  
 DB 853 DYLVRKRTSSNSVYTYCTCTYERPIAFTRSKKLWQFNSNNGNSARQGVYVYTD- 912  
 QY 923 ---EDYCELEDIVADGRVLVASENQEILKQKLIKALFCLARPOKYFKYTAQSEPMF 979  
 DB 913 RLPEDYQULIEVADGRVLVASENQEILKQKLIKALFCLARPOKYFKYTAQSEPMF 972  
 QY 980 PRSFIRLRKSVRFRLRPYK 999  
 DB 973 PRSFIRLRKSVRFRLRPYK 992

RESCUT 5  
 DB 081X30 PRELIMINARY; PRT; 993 AA.

AC 081X30;  
 DT 01-MAR-2003 (Trembl) 23; Last sequence update;  
 DT 01-MAR-2003 (Trembl) 23; Last sequence update;  
 DE CUB and EGF containing protein.  
 GN CEGF3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Pfarr N., Bahr A., Cichutek A., Joebert R., Zabel B.U., Schmidt E.R.,  
 RA Hackel T., Winterpacht A.;  
 RA "Novel human gene family (CEGF) encoding mosaic proteins with EGF-  
 RT like, STR2 and a CUB module: cloning and expression analysis."  
 RL Submitted (NOV-2002) to the EMBL/Genbank/DBS databases.  
 DR EMBL; AF452494; AAT6828.1;  
 SQ SEQUENCE 993 AA; 109282 MW; 19BBE05627BEAF4 CR664;

Query Match 60.6%; Score 3397; DB 4; Length 993;  
 Best Local Similarity 58.3%; Pred. No. 4e-233;  
 Matches 596; Conservative 136; Mismatches 212; Indels 78; Gaps 9;

QY 22 LPPLLALAGAVPGRGAAAPQEDVDECAQGLDCHADALCQNTPTSYKSCPKPGYQSEG 81  
 DB 6 VPGCLIVLTVARAAGYSKAAQVDECEGTNDCHIDA-CQNTPTSYKSCPKPGYQSEG 65  
 QY 92 RQCEIDIECGNE-LNGCVHDCANI PGNYRCTCFDGFMLAHGHNLDVDECLENNNGGCH 141  
 DB 66 KQCKQVDECEIDIECGNE-LNGCVHDCANI PGNYRCTCFDGFMLAHGHNLDVDECLENNNGGCH 125  
 QY 142 TCNVVGSYECCEKEGFLSDNQHICHRSEGLSCMKKHCQSCHICKAPRGVACER 201  
 DB 126 TCNVVGSYECCEKEGFLSDNQHICHRSEGLSCMKKHCQSCHICKAPRGVACER 185  
 QY 202 PGFELAKQK 261  
 DB 186 PGFELAKQK 237  
 QY 262 VTESNTSVYSGDKKVRRLMETCAVANNNGCDBTCKDTSRGVHSCGVGFTLQDGTCK 321  
 DB 238 VTESNTSVYSGDKKVRRLMETCAVANNNGCDBTCKDTSRGVHSCGVGFTLQDGTCK 275  
 QY 322 KQIDECQTRNGGCHFCNKIVGSPDQCKGKPKLLTDEKSCQVDECSLDRTCDHSCIN 381  
 DB 276 KQIDECQTRNGGCHFCNKIVGSPDQCKGKPKLLTDEKSCQVDECSLDRTCDHSCIN 335  
 QY 382 PGTACACNRGTYLGGFTHCQDQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQK 440  
 DB 336 PGTACACNRGTYLGGFTHCQDQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQK 395

QY 441 DCEVKGKLLPTSVSPVSLHCGKSGGDCGFLRC----- 474  
 DB 396 DTEPLKCGSGASKAMLSNRSKDKDTCLTCSRARFLPESNGFTVSGCTPSPAA 455  
 QY 475 -----HSGIHLSSD-----VTTIRTSVTFKLNKCGK--SLKNALLPFEGLRALPERHS 521  
 DB 456 PARAGHNGNSTSNHCHAEAVLSIKORASFRIKDKAKRLHLRNKSKTEBAGRITSPGAP 515  
 QY 522 SVKESFRVNLTCSSGKQVPGAPRPRSTPKKFI-ITVEFEJFTNKXETATASCDLSCTIKR 586  
 DB 516 CSECQVTFHLKCDSSRKGRKRRATPPCKEVTRLTLLEAEVRAEETTAACGLPCCJQR 575  
 QY 581 TEKRLAKAIRTLRKAVHREOFHQLSGMNLJVAKKPPTSEDA---SSCGVQGHAEHQ 617  
 DB 576 MERRLKGSJKLRKINSINDRFLRLAGLDYELAHKPLGLVAGRAAPMESCRCQGRAGT 635  
 QY 638 CVSCRAQTYDGAERECCLCPNGTQNEBGMTCBPCEPRPCNSGALKTPPEAMNSECGGL 697  
 DB 636 CVSCPGQTYHGGQTEQVCPAGTFOERGGQLSCDLCEGSDAHGFL---GATNVTTCAQC 692  
 QY 698 CQPGESADGFAFQCLALGTQCPAGRTSCFPCGGG-LATKQGATSFQDDETRVQCSFG 757  
 DB 693 CQPGHVSQGFRCPCPCRGTYQCPAGRTLCFPCGGG-LTKHGAISFQDDETRVQCSFG 752  
 QY 758 HEYNTTHRCIRCPGTYQPEFGKNNVCSPGNTTDEPGSTNTQCKNRRCGSGELGFT 817  
 DB 753 HYNTSIRHCIRCMGSIQDPFRQNFCSRCSPKSTCTDQGSTSVACQCKRCCJELGFT 812  
 QY 819 GYIESPNYPGNYPANTCTWTINPPKRIILIVPEIFLPIEDCGDYLVWKRTSSSNV 877  
 DB 813 GYIESPNYPGNYPAGCEICWNIINPPKRIILIVPEIFLPSDECGDYLVWKRTSSSNV 872  
 QY 878 TTYEFCQTYEPIAFTSRKSLMIQPKSNEGNSAGFQVPTTYTTEDEYQELIEDIVRCGR 937  
 DB 873 TYEFCQTYEPIAFTSRKSLMIQPKSNEGNSAGFQVPTTYTTEDEYQELIEDIVRCGR 932  
 QY 938 LYASENHOELIKDKLILKFLVLAHPONYFYTQESREVPASFIILRLKSVRFUR 957  
 DB 933 LYASENHOELIKDKLILKFLVLAHPONYFYTQESREVPASFIILRLKSVRFUR 952  
 QY 998 YK 999  
 DB 992 YK 993

RESULT 6  
 Q9EQC6 PRELIMINARY: PRT: 961 AA.  
 AC Q9EQC6:  
 DT 01-MAR-2001 (TREMblrel). 16, Created)  
 DT 01-MAR-2001 (TREMblrel). 16, Last sequence update)  
 DE 01-MAR-2003 (TREMblrel). 23, Last annotation update)  
 DE EGF-related protein SCUB1.  
 GN SCUB1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_taxid=10090;  
 RN NCBI\_taxid=10090;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=3H1;  
 RX MEDLINE=20541712; PubMed=11087664;  
 RA Grimmond S., Lazder R., Van Hateren N., Siggers P., Hulsebos T.J.M.,  
 RA Arkell R., Greenfield A.,  
 RT Cloning, Mapping, and Expression Analysis of a Gene Encoding a Novel  
 RT Mammalian EGF-Related Protein (SCUB1).  
 RL Genomics 73:74-81(2000).  
 CC -1- SIMILARITY: CONTAINS 1 CUB DOMAIN.  
 DR EMBL: AF276425; ANGI25939.1; -.  
 DR HSSP: P00742; IHG.  
 DR MGI: 1890616; Scubel.  
 DR InterPro: IPR000152; Asx\_hydroxyl.

DR InterPro: IPR000859; CUB domain.  
 DR InterPro: IPR001881; EGF Ca.  
 DR InterPro: IPR006209; EGF like.  
 DR InterPro: IPR001491; Thrombomodulin.  
 DR Pfam: PF00431; CUB; 1.  
 DR Pfam: PF00008; EGF; 9.  
 DR PRINTS: PR00507; THROMBOMODULN.  
 DR SMART: SMC0042; CUB; 1.  
 DR SMART: SMC0179; EGF CA; 6.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 6.  
 DR PROSITE: PS01180; CUB; 1.  
 DR PROSITE: PS01186; EGF\_2; 6.  
 DR PROSITE: PS01187; EGF CA; 6.  
 DR ECF-like domain.  
 KW SEQUENCE 961 AA; 103989 MW; 966726C4E3BBD33E8 CRC64;  
 Query Match 58.21; Score 3262.5; DB 11; Length 961;  
 Best Local Similarity 60.91; Pident No. 1,3e-281;  
 Matches 584; Conservative 99; Mismatches 181; Indels 95; Gaps 14;  
 QY 10 GAA---WAVLLJLLLPRLLLAGAVPRGGRAGPQEVCECAQGLDCHADALQNTP 66  
 DB 2 GAAVVRHLSLLALGARQQLVQSGLPG-----AVDVECSBGTDCHDAICQNP 54  
 QY 67 TSYKSCSPGYGSGRQCEIDECQNE-LNGGCVHCLNIPGNYRCTCFDGFYLAHDGN 125  
 DB 55 KSYKCLKPGYKSGRQCEIDECQNE-LNGGCVHCLNIPGNYRCTCFDGFYLAHDGN 114  
 QY 126 CLDYDECLNNGGQQTCTVNVMSYECCKEGFLSDNQHTCHIREEGLSCKMKDHGS 185  
 DB 115 CLDYDECLNNGGQQTCTVNVMSYECCKEGFLSDNQHTCHIREEGLSCKMKDHGS 174  
 QY 186 HICKARSGVACCEPRGFEFLANQRDCLTCYHNGSCGHSQDPTADGEGSCHPQYM 245  
 DB 175 HICKETKGVACDPRGFEFLANQRDCLTCYHNGSCGHSQDPTADGEGSCHPQYM 234  
 QY 246 HTDGRSCLREDTVLEVTESNTSVVDGKRVKRLMETCAVNGGCDRTCKDTSGVH 305  
 DB 235 HADGRTCI-----ETCAVNGGCDRTCKDTAGVR 264  
 QY 306 CSCPVGFTLQDGTCKDICEQTRNGGCHFKNIYGSFDCGCKKFKLLTDEKSCQDV 365  
 DB 265 CSCPVGFTLQDGTCKDICEQTRNGGCHFKNIYGSFDCGCKKFKLLTDEKSCQDV 324  
 QY 366 DECSJERTCHSCINHGTPACACNRYTLYKSTHCDTNECSINNGCCQVCNTVGSY 425  
 DB 325 DECSJERTCHSCINHGTPACACNRYTLYKSTHCDTNECSINNGCCQVCNTVGSY 384  
 QY 426 ECGCHPGYKJHMKKDCVEYKGLPTS-VSPVSLHCGKSGGDCGFLRC----- 474  
 DB 385 ECVPRPGRRLHMKDCVEMNGCLSRKSAQAOLSGKVGIVENCFLGSLFMPDS 444  
 QY 475 HSGIHLSSDY-----TTIRTSVTFKLNKCGSL--KNA 505  
 DB 445 ESSYILSCGVPLQGLKTLPRNGTSSSTGPGCSDAFTPIRQARFKIRDAKHLQPSQ 504  
 QY 506 ELPFEGGR-PALPERKHSVSESFRVYVNLTCSSGKQVPGAPRSTPREM-FITVEFELET 563  
 DB 505 ERANDTLRAPLQNCVHT-----FVTLKDDSKK--RRGRKSPSKVEHITAEFEVEM 556  
 QY 564 NQKVTASCCJSTIVKTEKRLKRAITRLKAVHREOFHQLSGMNLJVAKKPPTSERQ 623  
 DB 557 KVDBASGTCENQCKRKAQESLQALITLKSIGRNOFYQVLTREVQRPKALEGT 616  
 QY 624 AESCGVQGHAEHQCVSCRAQTYDGAERECCLCPNGTQNEBGMTCBPCEPRPCNSGAL 683  
 DB 617 G-TCGIGC-LQDGKCVCAAGTYFSGDPGCGPCVSGTQVDMESQJLCTPCP---SSSG 672  
 QY 684 KTPFAMKSCFCGSLCGQGEYSADGFAFQCLALGTQCPAGRTSCFPCGGG-LATKQGAT 743  
 DB 673 GLAGARVSSCGQCGSGYFSAAGFRCQACPVTGYQPEPRTQCFPCGGGLTKHGTGA 732  
 QY 744 SFQDCEFRVQCSFGHFTYTTTHRCIRCPGTYQPEFGKNNVCSPGNTTDEPGSTNITQ 803

```

Db      733 SFQDEAKVHCSPGHGHYNTTHRCIRCPVGYQPFQGNHCISCNGNSTLDFJ3STVIAH 792
Qy      804 CKNRRGGELGDFGTGYIESPNYPGVNPANTECTWTINPPKRRIILIVPEFLFIEDDGC 863
Db      793 CQNGHGGELGDTGYTGYIESPNYPGVNPANAEVWIAAPPKRRIILIVPEFLFIEDDGC 852
Qy      864 DYLVNMRKTSSNSVTTTGTCTYERPAFTSRKSKLMIQFKSNEGNSARGCVPIYTD 922
Db      853 DYLVNMRKSPSTSVTTGTCTYERPAFTSRKSKLMIQFKSNEGNSARGCVPIYTD 911

RESULT 7
Q8NAU9 PRELIMINARY; PRT: 880 AA.
AC      38NAU9;
DT      01-OCT-2002 (TEMBLrel. 22, Created)
DT      01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DT      01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE      Hypothetical protein FLJ34743.
OS      Homo sapiens (human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC      Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
CX      NCB:Taxid=9606;
RN      [.]_Taxid=9606;
RP      SEQUENCE FROM N.A.
RA      Nimotoya K., Magatsuna M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA      Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA      Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA      Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isoro Y.,
RA      Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA      Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa A.,
RA      Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA      Suzuki Y., Sugano S., Nagahara K., Masuno Y., Nagai K., Isogai T.;
RT      "NEDO human cDNA sequencing project.";
RL      Submitted (Jul-2002) to the EMBL/Genbank/DBJ databases.
CC      - - - SIMILARITY: CONTAINS 1 CUB DOMAIN.
DR      EXBL:AKC92062; BAC03798.1;
DR      InterPro: IPR000152; Asx_Hydroxyl.
DR      InterPro: IPR000859; CUB_domain.
DR      InterPro: IPR001881; EGF_Ca.
DR      InterPro: IPR006209; EGF_like.
DR      InterPro: IPR006210; EGF.
DR      InterPro: IPR001481; Thrombomodulin.
DR      Pfam: PF00431; CUB; 1.
DR      Pfam: PF00008; EGF; 5.
DR      PRINTS: PR00907; THROMBOMODULN.
DR      SMART: SMC0042; CUB; 1.
DR      SMART: SMC0181; EGF; 5.
DR      SMART: SMC0179; EGF_Ca; 5.
DR      PROSITE: PS00010; ASX_HYDROXYL; 3.
DR      PROSITE: PS01180; CUB; 1.
DR      PROSITE: PS01186; EGF_2; 5.
DR      PROSITE: PS01187; EGF_Ca; 3.
KW      Hypothetical protein.
SQ      SEQUENCE 880 AA; 97239 MW; F0BDEE85266FD60 CR664;
Query Match 52.0%; Score 2916; DB 4; Length 880;
Best Local Similarity 57.4%; Pred. No. 9e-251;
Matches 518; Conservative 127; Mismatches 186; Indels 72; Gaps 10;
Qy      146 WNGSYECCKEGEFSDNHTCIHSEBGLSCMNKDHGSHI CKAPRGSVACGECR2SFE 205
Db      1 WNGSYECCKEGEFSLDNHTCIHSEBGLSCMNKDHGSHI CKRET PKGGLACGCRPGFE 60
Qy      206 LAKNRDCLITCNHNGGCGQHS CDDTADSPFECSCHPQYCKHTDGRSCLEREDTVLEVTES 265
Db      6 LTKNGRDCLITCNVNGGCGQH CDDTDEGPRGCHIKFVLTGDKTCTI----- 108
Qy      266 NNTSVVGDGKRYKRL-----LMETCAVNGGCDRTCKDTSTGVHSCPVGFTLLODKRT 320
Db      109 -----GERRLECHIPTQAVSNETCAVNNGGCDKSCDAAAGVHCTCPVSGFMLOPDKRT 161

```

```

Qy      321 CKDIDECQTRNGGCHFGKAVIGSFDDCGKRSFKLLTDEKSCQDVDECSLDRTCRHSCIN 380
Db      162 CKDIDECRLNNGGCHICNITVGSFECSSKKYKLLINERNCQDIDESCFFRTCHICVN 221
Qy      381 HGTGACACNRYATLYGFTGCDTDECSINNGCCQGVNCTVGSYECCHPZY-ELHWNK 439
Db      222 TPSPGCLCHRGYLLGITHGCDVDECSINRGGCFPGINTGSIQCTCPAGCGALHNG 281
Qy      443 KDCVEYKGLPTVSERVSILHCGKGGGCGELRC----- 474
Db      282 KDCTEPLKQSGSGASKAMVSCNRGSKTKTALCTPSRARFLPESNGPTVAGTSPRA 341
Qy      475 -----HSGIHLSSD-----VTRITSTFELNKGK--S-KNAELPREGARPA:PEH 520
Db      342 APARAGHNGNSTNSNHCHEAAVLISIKRSPFKIKDAKGLHNNKXTEAGRLTGPGA 401
Qy      521 SSVKSEFRVYNITGSSGKQVPVAPGPPSPFKMPF--TVFELETKNKEVTASCDLSCIVK 579
Db      402 PCSECGVTFIHLKQDSSRKGRARRTPGKEVTRLTLELAEVAAEFTTACGCLPCLRO 461
Qy      580 RTEKRLKAKIRTLKAAVHREOFTLQSGNMLDVAKKPPRTSERQA---ESGVGGQHAEN 636
Db      462 RMERRLKSGSLKMLRKSLNODRFLPLAGLDVLAHKGVLVAGERAEPMESCRPGHRAQT 521
Qy      637 GCVSCRAGTYVCGARECILLCPNGTFQNEGQMTCEPCRPNGSGALXTPEAMNMSECG 696
Db      522 KCVSGQGYVHGOETEQVCPGAGTFOEREGQISCDLGGSDAHGL---GATNVITCG 578
Qy      697 LCPGSEYSDGFAPQLCALGTFQPEAGRTSCFPGCGSLATHQGATSFQDCETRVQCSF 756
Db      579 QCFPGHSDGFPKQPCPRGRGTYCPAGSTLCPGCGGLTTHGEGALSFQDCLTKVQCSF 638
Qy      757 GHFVNTTHACIRCPVGYQPFQGNKVCSCFQNTTDFDSTNTTQCKNRRGGELGDF 816
Db      639 GHYVNTSHRCIRCMGYSYCPDPFRQNFCSRCPGNTSTDFDGSSTVAQCKNRRGGELGDF 698
Qy      817 TGYTESPNYPGVNPANTECTWTINPPKRRIILIVPEFLFIEDDGCGLVNRKSSNS 876
Db      699 TGYTESPNYPGVNPANTECTWTINPPKRRIILIVPEFLFIEDDGCGLVNRKSSNS 758
Qy      877 VTTGTCTYERPAFTSRKSKLMIQFKSNEGNSARGCVPIYTDYEDYCELIJEDIVRG 936
Db      759 ITTGTCTYERPAFTSRKSKLMIQFKSNEGNSARGCVPIYTDYEDYCELIJEDIVRG 818
Qy      937 RLYASENHQTILKCKLILADVLAHQNYRYTQOERENVPKRFILLSKYSRFLR 996
Db      819 RLYASENHQTILKCKLILADVLAHQNYRYTQOERENVPKRFILLSKYSRFLR 877
Qy      997 PYK 999
Db      978 PYK 880

RESULT 2
Q8NAU9 PRELIMINARY; PRT: 581 AA.
AC      38NAU9;
DT      01-OCT-2002 (TEMBLrel. 22, Created)
DT      01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DT      01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE      Hypothetical protein FLJ34691.
OS      Homo sapiens (human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC      Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
CX      NCB:Taxid=9606;
RN      [.]_Taxid=9606;
RP      SEQUENCE FROM N.A.
RA      Tashiro H., Yamazaki Y., Watarabe K., Kumagai A., Itakura S.,
RA      Fukuzumi Y., Fujitani Y., Komiyama M., Sugiyama T., Irie R.,
RA      Otsuki T., Sato K., Wakamatsu A., Ishii S., Yamamoto J., Isoro Y.,
RA      Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA      Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Magatsuna M.,
RA      Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,

```



```

QY 509 ----PEG-----LPALEEKHS--VNESF- 527
DB 610 KCDPSSGTGEKCEQICRNGYWGVCAMKSCKLDEPSTGSCDEDEKSCDDGCPGDFY 669
QY 528 -----RYNLTGSSGKQVPGAPRSPTPKEMFIVVEFELNKEVNASDLSIVKATE 592
DB 670 GSQCNLKCRCMDCPNGRCDF-VFGYCTCPDGLY-----GQSCKEKCFHFTFG 714
QY 583 KRLKRAIRTLRK-----AVHREQF-----HQL 605
DB 715 KRCRFPCKCARBNESGDEITGCRCKRPGYTHHCKMKSCGSLFGAGCAMKSCPRGIR 774
QY 606 SGMLDVAKRPPR-----TSEQAESCQVGGHAENQC----- 638
DB 775 DPTVGDCCTKCPAGYQGNLCPGCPAGYFGYDCEQKSCADVASPHKSKYCHVGTCTC 834
QY 639 -----VSCAGTYDDAGERCLCPNGTQNE-EGQMTCEP-----CP- 675
DB 835 LPKXTGLCDQSCAPNTYGPNCATTC-SCVNAKACDSDSGCHCTPGFYATGSEVCPFG 893
QY 676 RFG-----NSGALKTPEAMNMSGCGGLCPGFEYSADGFAPCQLCALGT 718
DB 894 RFGIDCMQLCKCQNGALCDTSGSCGAPGWSGKKCDKACAPG-FGADCSKKCD-CADGM 952
QY 719 F-QPEAGRTSCFP-----CGGGL-----ATKQGAISFG---DCETRV--- 752
DB 953 HCDPSDGEICFPKPKGHKCDCTCDGLFGAGCKGICSCQNGATCDVTSGCCRPGRWG 1012
QY 753 -----QCSPGHF-----YNTTTHRC-----IRCPVGTIOPEFGKN 783
DB 1013 KCDRPPCDGFRFGECNAICCTTNTDNTSMNPPVACDHTGSCRCRCPAGTGPD----- 1067
QY 784 C-VSPGNTTDPDSTNTITQCKNRRCGELGDFTYIESPNYGNYPANTEC---TW11 839
DB 1068 CQTSQCP--LGRHGECRHSQCQSN---GASCDRVLTGFDCCPSGFMGNKCSCECEGLMGS 1122
QY 840 N 843
DB 1123 N 1123

```

RESULT 10  
088281  
AC 088281; PRELIMINARY; PRT: 1574 AA.  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE ME6F6.  
CN ME6F6.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCB\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sprague-Dawley; TISSUE=Brain;  
RX XEDLINE=98360089; PubMed=9693030;  
RA Nakayama N., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.,  
RT "Identification of high-molecular-weight proteins with multiple EGF-  
RT like motifs by motif-trap screening";  
RU Genomics 51:27-34(1998)  
DR EXBL: AB011532; BAA32462.1; -.  
DR HSSP: P00736; IAPQ.  
DR InterPro: IPR00152; Asx\_hydroxyl.  
DR InterPro: IPR001881; EGF\_Ca.  
DR InterPro: IPR006209; EGF\_like.  
DR InterPro: IPR002049; Laminin\_EGF.  
DR Pfam: PF00008; EGF\_24.  
DR PRINTS: PRO0011; EGF\_LAMININ.  
DR SMART: SM00179; EGF\_CA; 4.  
DR PROSITE: PS00010; ASX\_HYDROXYL; 5.  
DR PROSITE: PS00022; EGF\_1; 23.

```

DR PROSITE: PS01186; EGF_2; 23.  
DR PROSITE: PS01187; EGF_CA; 5.  
KW EGF-like domain.  
SQ SEQUENCE 1574 AA; 165445 MW; 2B40531DE8F7FE27 CRC64;  
Query Match 15.0%; Score 841.5; DB 11; Length 1574;  
Best local similarity 26.3%; Pred. No. 1,76-65;  
Matches 245; Conservative 74; Mismatches 289; Indels 323; Gaps 40;  
QY 34 PGRGRAGPC---EDVDECAQGLDPCNADA-CQNPPTSYKSCSGPY--QSGRCQCEID 88  
DB 109 PGMGKPPQGEBCCSVDCAANAGCC--EGECNTVGGFYRCRPGVLQDQDKTCQVD 166  
QY 89 RCGNELNGGCVHDCINIPNYRCTCFDSEFLAHQDNCLVDECLENNAGCQHTCVN-V- 147  
DB 167 EC-RAHNGGQCRVNTPGSYLCEKPGFRJHTDRTCLAISSCTLGGCCQHCQVQLTV 225  
QY 148 GSECCCKEGFTLSNQHTCIHRSREG-SCKMKHGHGHIKEAPRGSVACSRPFELA 207  
DB 226 TQHRQCPRQYQLOEDGRVVRSS---PCAEGNGGCMHICQEL-RGLHGGCHFGYGLA 280  
QY 208 KQRPC--ILTGNHGSCQSCDDTDADPSCSCHPOYKMTDGRSCLERBDTVLEVES 265  
DB 281 ADKTCEDVDGALGACAHGCAHCCTGSGFKCVGAGYELGADSRQYRIE--MEIVNS 337  
QY 266 NTTSVVDGKRYKRLMETAVANNGGCDRTCKJSTGVHSCRPVFTLQDKTCKID 325  
DB 338 -----CEANGGSGSHGSHSTSTPLCTCPRGYELDEQKTCIDID 377  
QY 326 ECGTRNGGCHFECKNIVGFCGCKKGRKLTDEKSCQDVDECSLDR-TCQHSCTNPQT 384  
DB 378 DC-ANSPCCQACANTPGGYBCSFAGRRLNTDGGCEDVDCAAGHGGCHRSNLAGS 436  
QY 385 FACACNRGYTL-----YGFTH 400  
DB 437 FQCFCEAGRYLDEDRGCTSLSESVYDLDGR-LFVAP-LPH-LAVLDELPLRQPDYGAEE 496  
QY 401 -----CGD---TNEGSI-----NNGS-----CQ 415  
DB 497 EAAAEALRGENTLTERKPYCLDHSFCHDSCLTCDRCNGCTCFPGQDGDCCPEGMTGII 556  
QY 416 QVC-VNTVG---SYECQCHPGYKLMNKKDCEVEYKGL--LFTSVSPVSLH---CGKSG 465  
DB 557 ETCPPDTEGKNCSSPCTCQ-----NGCTCPVLGACRCRPGVS---GAREDDCCPGF 606  
QY 466 GDCGFLRCHSSILHSEVTTIRTSVTFKLENGK-SLXNAELPREG-LRPALEFGHSSVK 524  
DB 607 YGKHKRKKCHCA-----NRGRCHRLYGALCLCPGCG-YG----- 638  
QY 525 ESEFVYNLTGSSGKQVPGAPRSPTPKEMFIVVEFELNKEVNASDLSIVKATE 584  
DB 639 ---RFOHLACPFMAFGP-----TSEDCLC----- 660  
QY 585 LKRAIRTLRKAVHREDFHLQLSGXNLDVAKKPPFTSEQAESCQVGGHAENQC-VSCRA 643  
DB 661 -----EOSH-----TRCNPKDGSCSKAGQGBRCQAEES 692  
QY 644 GYYVGAREGCTL-----CENGT-ONEEGMTG- 671  
DB 693 GFGPGCRHRTCGPVAADPVSGECRTCCPGYQGEDGCGCPGTFGWNSSGSCSG 752  
QY 672 EPCPR-----PNSGALKTPKAMNMSGCGGLCPGFEYSADGFAPCQLCALGT-TPPE 722  
DB 753 APCHRVTGCLCPPEKGTG-----EDCAGDCPEGRMGLGCCEITPACGHSASCP 802  
QY 723 AGRTSCFPGGGLATKQCATSF---QDCETRVQCSGPHFYNTTHRCI----- 768  
DB 803 TGTGCLLF--GVGSRQDPTCSAGWVGCGQIRCAADGHCDDPTTGRCSCAPMTGJSC 860  
QY 769 ---RCPVGTIOPE-----FGKNCVSCPG 789  
DB 861 CRACDSGHGAPDCIHPNCNSAGHGCAVSG 891

```







OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99350211; PubMed=1043698;  
 RA Yang Q., Oca K., Tian Y., Kumar A., Wada J., Kashinara N., Wallner E.,  
 "Cloning of rat fibrillin-2 cDNA and its role in branching  
 morphogenesis of embryonic lung."  
 RT Dev. Biol. 212:229-242(1999).  
 RL EMBL: AF15060; AAC34439.1; ..  
 DR HSP; 23555; 1EXN  
 DR Interpro: IPR002086; Aldhyde\_dehydr.  
 DR Interpro: IPR000152; Asx\_hydroxyl.  
 DR Interpro: IPR001881; EGF\_Ca.  
 DR Interpro: IPR001438; EGF\_11.  
 DR Interpro: IPR006209; EGF\_1like.  
 DR Interpro: IPR002212; Fibrin-assoc.  
 DR Pfam: PF00683; TB; 9.  
 DR Pfam: PF00683; TB; 9.  
 DR PRINTS: SM00179; EGF\_CA; 42.  
 DR SMART: SM00179; EGF\_CA; 42.  
 DR PROSITE: PS00073; ALDEHYDE\_DEHYDR\_CYS; 1.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 43.  
 DR PROSITE: PS00022; EGF\_1; 2.  
 DR PROSITE: PS01186; EGF\_2; 36.  
 DR PROSITE: PS01187; EGF\_CA; 43.  
 KW EGF-like domain.  
 SQ SEQUENCE 2906 AA; 313371 MW; 9EB64E727044FE58 CRC64;

Query Match 12.8%; Score 720; DB 1; Length 2906;  
 Best Local Similarity 26.9%; Pred. No. 2,7e-54;  
 Matches 256; Conservative 96; Mismatches 308; Indels 292; Gaps 54;

QY 45 DVDECAQGLDCHADALCONTPSYKSCSPGYQ--ESQCEIDDEGNEINSGCTFC 102  
 DB 1192 DINECSLS-DNLCRNKCKVMGTYQSCNPQYQTPROGCSDDICE-MKNGSDDTQC 1249  
 QY 103 LNIQNVYRCTCFDGSFLADHGNCLDVECLENN-----GGQHTCVVWVSYECCCKEG 157  
 DB 1250 TISEGSEYECSCSEGALMPDGRSCADIDEC-ENNDIDICG--QCTNIPGEYRLCTIDG 1305  
 QY 158 PFLSDNHTCHIRSEEGSCMKKHGCHICK---EAPRGSVACERPGFELAKNGRDC 213  
 DB 1306 FMASVDMKTCIDVNECDLN-----PNCMFGECEHTKGSFICHQDQSVKKGATGC 1357  
 QY 214 --ILTCNHNGGCG--QHSCDDTADGPESCHQYKMHNDGRSCLERETVLEVESNTTS 269  
 DB 1358 TDVDECEIGAHCNDHNASCLNVPSFKSCREGKV--GNGIKCIDLDJSC-----ANGT- 1408  
 QY 270 VVDGDKRYRRLMETCAVNNGGCDRTCKDTSTGVHSGCPVGFLLQDKTKCKDIDECQT 329  
 DB 1409 -----HCGSIN-----AQCNTPGSYRCACSEFT--GGFTCSVDECAE 1447  
 QY 330 RINGGCDH-FCKNIVSGPDCGCKGFKLLTDEKSCODVDECSLERTCDH-SCINHPGTAC 387  
 DB 1448 NINLCENGGCLNVPAVRECEMGFTPASDSRSQDIDECSCQNICVEGTGNNIPGMFHC 1567  
 QY 388 ACHRGYTL-YGFTHGDNMGCS--IN--NGGCGVCVNTVSGYECQCHPGYKLLHNNKDC 442  
 DB 1508 ICDDGGLDRTGCHCTDIDECADPNCVNG---LCVTPGRIEKNCPDPDQLVATGVGC 1563  
 QY 443 VEVR-GLLPTSVSPVSLHCGKSGGDDG---CFLCHSGIHLSSDVTIIRSVTFKLNKG 498  
 DB 1564 VDNRVGNCLYKRGPR-----GDGSLCKTEWGVGVSSSSCC----- 1599  
 QY 499 KSLKNAELFPGLPALPBRKSSVKSFRYNNLTSSSKCYPAAGPS-PRKEXFTVE 558  
 DB 1600 -CSJGKMGNPCECTCPV-----NSTEYYSL-----CPGEGG--FRPNQITILE 1641  
 QY 559 -----FELETNQ-----KEVTASCO----- 573  
 DB 1642 DIDECCLPGLCGGNCINTGSGFCQCEPGYLLSEETRICGIDIDECFAHRAVCGPGTGY 1701

QY 574 -----TSCIVKTEKRLRAIRTLFKAVERQFHLQSGMN--LDVAKK----- 615  
 DB 1702 NTLANTTICLPP-----YMGVNGHNMOMDRKFCVRYNGTTC 1741  
 QY 616 ---PRTERQAESGVGQGAENQ-CVSC-RAGTY----- 646  
 DB 1742 ENELPNVTVKRC-CCYVVGKAMNRCFCPPGADFTICGNIIGFTFDIHGTCAVD 1800  
 QY 647 YDGAERCLCPVGTQNEGGVTCPPCRPNSSGLKTPKAN-----XSEGG--- 695  
 DB 1801 IDCKEIPGICANGVCINQGFRC-CP-----TGSFYNDLLVCEDIDECSDND 1850  
 QY 696 GLCY-----PGEVSADCFAPQCALG-TFQPAQ--RTSCFP-----COGLATKH 739  
 DB 1851 NLCORVADCCNEPGRFCE-----CAAGFKSLPAGACVDRHCELEINVCSSGLCVL 1903  
 QY 740 CGATSFQ-DETRVQCSPGHFYNTTHRCIR--CPVGYQPEFGKNCVSCPG-----N 790  
 DB 1904 CG--SYQCIENKGFASQCTMCKEYDECEHPCGNGTCKNTVGSYNCLCYPPFELTHNN 1961  
 QY 791 TTDFDGSNTI--TCKNRCGSELGDF-----TGYTESPNYGNVPAATFC 835  
 DB 1962 DCLDIDECSSFFGQVCRNRCFENEISGFKCLNBSYEELPCGXNCTDTN-EC 2012

## RESULT 14

966CPA  
 ID Q96JPR PRELIMINARY; PRT; 2809 AA.

AC Q96JPR; 01-DEC-2001 (TREMBlrel. 19, Created);  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update);  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update);  
 DE Fibrillin3.  
 CS KIAA1776.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteiostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=21245130; PubMed=1147906;  
 RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XX.  
 The complete sequences of 103 new cDNA clones from brain which code  
 for large proteins in vitro.";  
 RL DNA Res. 8:85-95(2001).  
 DR EMBL: AB053450; BAB47408.1; ..  
 DR Interpro: IPR00152; Asx\_hydroxyl.  
 DR Interpro: IPR001881; EGF\_Ca.  
 DR Interpro: IPR006209; EGF\_1like.  
 DR Interpro: IPR002212; Fibrin-assoc.  
 DR Pfam: PF00683; TB; 9.  
 DR Pfam: PF00683; TB; 9.  
 DR SMART: SM00179; EGF\_CA; 41.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 41.  
 DR PROSITE: PS00022; EGF\_1; 2.  
 DR PROSITE: PS01186; EGF\_2; 36.  
 DR PROSITE: PS01187; EGF\_CA; 40.  
 KW EGF-like domain.  
 SQ SEQUENCE 2809 AA; 300323 MW; 20C04CC006C0161F CRC64;

Query Match 12.7%; Score 710.5; DB 4; Length 2809;  
 Best Local Similarity 24.8%; Pred. No. 1.8e-53;  
 Matches 262; Conservative 79; Mismatches 316; Indels 401; Gaps 50;

QY 45 DVDECAQGLDCHADALCONTPSYKSCSPGYQ--GEGRCQEDIDECG----- 91  
 DB 1071 DVDECARPPLCPG-CTGNTDGSYKCGPQHEJAKTACEDIDECSLSDGLCPHSGC 1129  
 QY 92 -----NCT-----NGGCHDCLNIPGNVYRCTCPDGFYLAH 121

```

Db      1130 VNVIGAFCCSHAGFQSTPRQGCVDINECRVQNGCCDHCINTESSYRCSGCGYSLNF 1189
Cy      122 DGHNCLDVDECLENNNGG--CHTCVNVWMSYECCCKEGFLSDNCHTCHHSEBGLSCMK 180
Db      1190 DGRACADVDCEBENRVCDOGHCHTMNGHRCLODYDFMATTDMRRCVDDVE-----CDLN 1245
Cy      181 DHGSHITKEAPRGSVACEPRPFELAKNORC--ITCHNGGCGCH--SCDCTADGPE 236
Db      1246 PHICLHGDCEHTKSPFCHCQLGYVYKRGATGCDVDECEGCHNCDSHASCINTIPGFS 1305
Cy      237 CSHPOYKMTDGRSCEREDTVLETSNNTSVVDDCKRKRLLMETCAVNNNGGDR 296
Db      1306 CRLLPGWV--GDGFEBCHLDDEICISOEHRCSPR-----GD----- 1337
Cy      297 CKDTSTGVHSCSPVGFLLDLGKTKCIDECQTRNGGCH--FCCKIVSFDGCKKGFKL 355
Db      1338 CLNVPGSYRCTCRGGA--GDGFCEBDECAENVLDLDMGCGLNAGRGVCEGEMGDFP 1395
Cy      356 LTDEKSCDVDECSLDRTCH--SCINHPGTACACNRYTL--YGFTHCGDTNCS--IN- 410
Db      1396 TEDHACQDVDECAQGNLCARFSCENLPGRFCINQGYELDRGGGCTDINECADPVNC 1455
Cy      411 -NGGCGQVCVNTVSYEGCGHGYKLMNKKDCVEVGLPLTSVSPVSLHCGSGGDDG 469
Db      1456 ING-----VCINTPSSYLCSGQDFELNBSGVGVDT-----AGN 1491
Cy      470 CFLRCH-----SGIHLSD--VTTRTSYTFKLNKGKSLKNAELFPBGLPPLPEKHSV 523
Db      1492 CFLETHRGGDSGISCSAIGVGTBASCC-----CSLGRAMGNPELCP-----M 1536
Cy      524 KESRRYVNLTCSSGQKQVGAEG--RPSTPKEMFTVEFELENNQEVNASCO--LSCTYKRT 581
Db      1537 ANTEYERL-----CPGEGGFQPNR-----ITVLEIDIECQELPGLCQSGSPCV--- 1580
Cy      582 EKRLKAIKRLKAVHROPHQLSGNMLDVAKKP--RTSE-----RQASCGVGGGA 634
Db      1581 -----NTBSPFCCEPFGYHLSHNRICEDTDEGTHSG-- 1614
Cy      635 ENQVSCRAGTYDGAERECILCP-----NGTFQNE----- 665
Db      1615 -----ICPGTCVNTLGNVTCVPAEY--QVANGNNCMOMKSCVCFRRVNGTCQHELAFNV 1663
Cy      666 -----EGQ--MTCEPQPRP-----GNSEA--LKTPEAMNYSGG-- 695
Db      1670 TRKMCCSYNIGQANRPCEACPPIPSDYOILCGNQAPGFCTIHTGKPLDIECGEIP 1723
Cy      696 -----GLC--QPEYSAD--GF-----APCQ----- 712
Db      1730 AICANGICINQISFRCECPAGFVNNSTLLACEDVDEGCSHESPCCQADCNTPGSYRC 1783
Cy      713 -----JCLGTPQPEAGRTSCF-- 723
Db      1790 KOTRSYKLSPGAGACVQNECREIPNVCSHGQMDTEGSMYLCIRHG--FQASADTLCMDI 1843
Cy      720 -----PCGGG-----LATRQAGATSPQCCETRV--CGSPGFHYTT- 763
Db      1849 DECDROPFCGNGTCKNIISVYCLCPFGVVTNMGCVDPDCTTLVGQVCRFGGLATAG 1968
Cy      764 -----THRCI--RCPVGYOPEFGKNNCVSCPKNTTTFDGS 798
Db      1909 SFHGLCQCGFELTADKKACVDTNECLSLAGTCLPOTCONLBSFRCT--CPGPGFVGCFHC 1967
Cy      799 TNLTCCKNR-----RGGELGDF-----TGYESRN 824
Db      1968 IDIDECSEEPNLCFTCTNSPGSFQCLCPGFSVLSDR 2005

```

RESULT 15

C9C435

ID 08C435

PRELIMINARY:

PRT:

810 AA.

AC

01-MAR-2003 (TREMblrel: 23, Created)

DT 01-MAR-2003 (TREMblrel: 23, Last sequence update)

```

DT      01-MAR-2003 (TREMblrel: 23, Last annotation update)
DE      Matilin 2.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX      NCBI_TaxId=10090;
RN      (1)
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Hippocampus;
RX      MEDLINE=2215463; PubMed=12466851;
RA      The FANTOM Consortium.
RA      The RIKEN Genome Exploration Research Group Phase 1 & II Team;
RT      "Analysis of the mouse transcriptome based on functional annotation of
RU      Nature 420:563-573 (2002).
RU      Nature 420:563-573 (2002).
CR      EMBL: AK081151; BAC18785.1;
SQ      SEQUENCE 810 AA; 90072 MW; 12CB0D391B6C27BC CNG64;

```

```

Query Match      12.4%; Score 695; DB 1.; Length 810;
Best local Similarity 30.0%; Pred. No. 8e-53;
Matches 160; Conservative 71; Mismatches 224; Indels 78; Gaps 14;

Cy      62 CQKPTSYKCGCKPGY--CGSAPCCED:DECGNELMGCCVHDCNIPGVNCTCFDGM 119
Db      253 CUNTSGSYLCKKQGGYIISTDKTCRIIDLQATE DHCCQGLCVNMLSSPVCQCGSYTT 311
Cy      120 AHGHNCLDVDECLENNNGGCHTCVNVWMSYECCCKEGFLSDNCHTCHHSEBGLSCMK 179
Db      312 AEDGRCAVDYCAEKNHGCHECVNASSYLCHREGFALNSCKTKSLD---YCAS 367
Cy      180 DHGSHITKEAPRGSVACEPRPFELAKNORC--ITCHNGGCGCHSCDCTADGPEG 237
Db      368 SHHGOHECVNA-QTSAQCLKGLMPEKTKRRIVYCAANKPGCHECVNTEGHYC 426
Cy      238 CSHPOYKMTDGRSCEREDTVLETSNNTSVVDDCKRKRLLMETCAVNNNGGDR 296
Db      427 KRGQVNLDPKTKTSRDRHCAQCHGEOLCLNTEESFVCCSSGFLINDLTKCSAD 486
Cy      285 TCANVNGGDRCKDSTGVHSCSPVGFLLDLGKTKCIDECQTRNGGCHFCNNIVGS 344
Db      487 YLLSNHGEYSVANTDKSPACQCEGHVLRDQKTCAXKLDSCALGDHSCHESSSDS 546
Cy      345 FCGCKKGRKLLTDEKSCQVDEEC--SLRTCDHSC--NHPGTACACNRYTL--YGFTHCG 402
Db      547 FVCCQPEGYILRDQKTCRRKDCVQCVNHHGSHLCVNSGSESYVCKLESPRLAEDGRCK 606
Cy      403 DTEGCSYNGSCQVCVNTVSYEGCGHGYKLMNKKDCVE----- 444
Db      607 RKNVCKSTQHGCEHNCVNNNSYLCRCEGFVLAEDGHCRCRTEGPIDLVFVIDGSKS 666
Cy      445 -----VKGLLFT--SVSPRVSLHCGSGGDDGCFLRHSGIHLSSDYTTTRTSY 492
Db      667 GEENFETVHFTVGLIDSLAVSPKA-----ARVGLLYSTQVRIEFT 708
Cy      493 FK-----LNSGKSLKNAELFPBGLPPLPEKHSVSKESFRVYVNLTCSSGKVP 541
Db      709 JGFFSANKMKCAVTHKMYKGSMTGLALGH-MERSEFTQVEGARPSQVVP 760

```

Search completed: October 22, 2003, 11:30:17  
 Job time : 130 secs